



Langevin Dynamics with Variable Coefficients and Nonconservative Forces: From Stationary States to Numerical Methods

Matthias Sachs, Benedict Leimkuhler, Vincent Danos

► To cite this version:

Matthias Sachs, Benedict Leimkuhler, Vincent Danos. Langevin Dynamics with Variable Coefficients and Nonconservative Forces: From Stationary States to Numerical Methods. *Entropy*, 2017, 19 (12), pp.647. 10.3390/e19120647 . hal-01976381

HAL Id: hal-01976381

<https://hal.science/hal-01976381>

Submitted on 10 Jan 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Article

Langevin Dynamics with Variable Coefficients and Nonconservative Forces: From Stationary States to Numerical Methods

Matthias Sachs ^{1,2}, Benedict Leimkuhler ^{2,*}  and Vincent Danos ³

¹ Department of Mathematics, Duke University, Durham, NC 27708, USA; msachs@math.duke.edu

² The School of Mathematics and the Maxwell Institute of Mathematical Sciences, James Clerk Maxwell Building, University of Edinburgh, Edinburgh EH9 3FD, UK

³ Département d'informatique, École Normale Supérieure, 45 rue d'Ulm, F-75230 Paris CEDEX 05, France; vincent.danos@gmail.com

* Correspondence: b.leimkuhler@ed.ac.uk

Received: 27 September 2017; Accepted: 22 November 2017; Published: 29 November 2017

Abstract: Langevin dynamics is a versatile stochastic model used in biology, chemistry, engineering, physics and computer science. Traditionally, in thermal equilibrium, one assumes (i) the forces are given as the gradient of a potential and (ii) a fluctuation-dissipation relation holds between stochastic and dissipative forces; these assumptions ensure that the system samples a prescribed invariant Gibbs-Boltzmann distribution for a specified target temperature. In this article, we relax these assumptions, incorporating variable friction and temperature parameters and allowing nonconservative force fields, for which the form of the stationary state is typically not known a priori. We examine theoretical issues such as stability of the steady state and ergodic properties, as well as practical aspects such as the design of numerical methods for stochastic particle models. Applications to nonequilibrium systems with thermal gradients and active particles are discussed.

Keywords: Langevin dynamics; fluctuation-dissipation theorems; nonequilibrium simulation; molecular dynamics; sampling; local thermal equilibrium; temperature gradients

1. Introduction

Langevin dynamics is a system of stochastic differential equations of the form

$$\frac{dq}{dt} = M^{-1}p, \quad (1)$$

$$\frac{dp}{dt} = F(q) - \Gamma(q)p + \Sigma(q)\dot{W}. \quad (2)$$

Here q, p represent vectors of the positions and momenta of the particles comprising together $2n \in \mathbb{N}$ degrees of freedom. The mass matrix M is symmetric positive definite. The force F is usually taken to be conservative, i.e., $F(q) = -\nabla_q U(q)$ for some potential energy function U . Langevin dynamics describes a physical system of particles moving under prescribed interaction forces and subject to collisions with particles of a “heat bath.” The friction matrix Γ , which is here allowed to vary with position, typically models drag on a set of distinguished particles due to the interactions with the surrounding environment, whereas the matrix Σ , also potentially varying with position, characterizes the stochastic effects of collisions. W represents a vector of independent and uncorrelated Wiener processes in \mathbb{R}^m , $W = [W_1, \dots, W_m]^T$, where the components of its formal time derivative \dot{W} are independent white noise processes such that

$$\mathbb{E} [\dot{W}_i(t) \dot{W}_j(t')] = \delta_{i,j} \delta(t - t'). \quad (3)$$

and

$$\mathbb{E} [\dot{W}_i(t)] = 0. \quad (4)$$

In case $\Gamma(q) \equiv \gamma \mathbf{I}$ where γ is a nonnegative scalar and constant, one may take $\Sigma(q) \equiv \sigma \mathbf{I}$, where $\sigma = \sqrt{2\gamma k_B T}$, T is the temperature and k_B is Boltzmann's constant. The resulting system can then be shown, under conditions reviewed in Section 2, to have a unique invariant distribution with density $\rho_\beta(q, p) \propto \exp(-\beta H(q, p))$, where $H(q, p) = p^T M^{-1} p / 2 + U(q)$ represents the total energy of the isolated deterministic system and $\beta^{-1} = k_B T$.

The emphasis in this article is on the general form (1) and (2), for which we will relax several typical assumptions. First, the system is not assumed to admit a universal fluctuation-dissipation relation, instead we assume only certain nondegeneracy conditions. Second, we do not assume a conservative force field. Such generalized forms of Langevin dynamics can be used to model diffusion (or thermal) gradients by particle simulation [1–3], as well as a variety of models for flocking [4–6], protein folding [7] and bacterial suspensions [8]. Problems with nonconservative forces are considered often in the physics literature, but the precise characterization of the convergence to stationary conditions is rarely discussed. In the general case, the form of the stationary distribution is non-obvious and the uniqueness of the steady state and its asymptotic stability are not well understood. An important contribution of this article is to show the existence of a steady state for the model (1) and (2) under relaxed conditions and to explicitly construct a Lyapunov function, thus allowing us to conclude the geometric convergence of observable averages.

Our approach to proving ergodicity is dependent on the fact that the friction and noise terms in the equations depend only on positions. On the other hand, within this class, and subject to the nondegeneracy condition on the friction and noise tensors, our treatment is general and has a considerable range of applications. Formally (1) and (2) includes dissipative particle dynamics (DPD), which is a momentum-conserving, 2nd order, gradient-type system [9,10], but we specifically exclude in this article cases for which Γ is not positive definite, which occur as a direct consequence of momentum conservation in DPD; ergodic properties for DPD systems in one dimension were discussed in previous work of Shardlow and Yan [11]. See also [12–14] for some special cases of Langevin-type systems with velocity dependent coefficients to which our formulation is not applicable.

To perform numerical discretization of the system (1) and (2), we proceed by splitting. This involves performing an additive decomposition of the stochastic vector field into several component parts, each of which can be exactly integrated (in the sense of distributions). The resulting stochastic maps can be composed to approximate the solution in a single timestep. Even for a particular choice of decomposition, there are many ways to combine the sequence of steps. The properties of various choices of the splitting method have been examined in detail for constant Γ and Σ in [15]. In that work it has been shown that the numerical methods constructed using certain components inherit the ergodicity properties of the SDE system, whereas the invariant measure of the numerical method differs from that of the SDE. The error in the invariant measure can be characterized in terms of the order of accuracy of averages of test functions. In general, it is found that symmetric compositions are preferred as they exhibit, under mild conditions, even order of accuracy for the invariant measure (thus one obtains second order accuracy for the same computational work as a first order scheme).

In the case of the more general systems considered in this article, the calculation of the Ornstein-Uhlenbeck solution, which is required at each step of our splitting-based numerical methods, becomes potentially demanding from a computational standpoint. In the methods of [15,16] this is done exactly, however in the current setting the solution of a Lyapunov equation would need to be obtained at each step in time, a calculation that would dominate the computational load in a large scale simulation. Therefore, we offer an efficient numerical procedure based on multiple timestepping [17], relying on a further splitting of the OU equation at each interior timestep.

In the last section of this article, we illustrate the theory in three numerical examples. The first is a particle system with an imposed diffusion gradient, as in [7,18–21]. Here the issue is to capture the correct density fluctuations as a function of position. In the second model, we incorporate a stirring force which alters the equilibrium state of a physical model. In the last example, we discuss the use of a blended Langevin model which draws on ideas from the literature of flocking, in particular the Cucker-Smale system. We use our theory to infer attractive steady states for the system, and characterize flocking tendencies by the use of two order parameters: one modelling the formation of *consensus* and the other characterizing the *peculiarity* which can be viewed as the average internal energy of isolated clumps of matter.

2. Stationary States of SDEs and Their Stability

In this section, we briefly outline the general theory on which our analysis of the ergodic properties of the system (1) and (2) in the next section is based. We are following in large parts the presentation in the review articles [22,23] and we refer the reader to these articles for a more detailed and comprehensive presentation.

Let $(x(t))_{t \geq 0}$ denote the solution of an Itô diffusion process of the form

$$\dot{x} = a(x) + B(x)\dot{W}, \quad x(0) \sim \mu_0, \quad (5)$$

taking values in a suitable domain. In this general presentation, $x(t) \in \Omega_x$ may represent either a position vector or else the combined vector of positions and momenta, i.e., $x(t) = (q(t), p(t)) \in \Omega_q \times \mathbb{R}^n$. In the latter case, we would typically either assume a compact configurational domain, as for example when periodic boundary conditions are used in the position space, i.e., $\Omega_q = L\mathbb{T}^n$, where $L > 0$ and $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ denotes the 1-torus, or else an unbounded domain, e.g., $\Omega_q = \mathbb{R}^n$. $(W(t))_{t \geq 0}$ represents a standard Wiener process in \mathbb{R}^m and the coefficients a, B are assumed to be smooth, i.e., $a \in C^\infty(\Omega_x, \mathbb{R}^{2n})$, and $B \in C^\infty(\Omega_x, \mathbb{R}^{2n \times m})$. The initial state of x is specified by the probability measure μ_0 , which throughout this article we assume to be such that $x(0)$ has finite mean and variance, i.e.,

$$\int_{\Omega_x} \|x\|_2^2 \mu_0(dx) < \infty.$$

2.1. The Associated Semigroup of Evolution Operators and Their Adjoints

The time evolution of the expectations of observables under the dynamics of the SDE is described by the semigroup of *evolution operators* $(\mathcal{P}_t)_{t \geq 0}$

$$(\mathcal{P}_t \varphi)(x) := \mathbb{E}[\varphi(x(t)) \mid x(0) = x],$$

for $\varphi \in \mathcal{S}, x \in \Omega_x$, where the expectation is taken with respect to the Wiener measure associated with the driving noise process $(W(t))_{t \geq 0}$ and $\mathcal{S} \subseteq \mathcal{M}(\Omega_x, \mathbb{R})$ denotes a set of *test functions* or *observables*, which is contained in $\mathcal{M}(\Omega_x, \mathbb{R})$, the set of all real valued measurable functions defined on the domain Ω_x . If the test function set \mathcal{S} is chosen appropriately, (e.g., $\mathcal{S} = \mathcal{C}_b^\infty(\Omega_x, \mathbb{R})$, where $\mathcal{C}_b^\infty(\Omega_x, \mathbb{R})$ denotes the set of all smooth bounded real valued functions defined on Ω_x) the action of \mathcal{P}_t corresponds to the solution of an initial value problem, i.e., $(\mathcal{P}_t \varphi)(x) = u(x, t)$, where u solves

$$\frac{\partial}{\partial t} u(x, t) = \mathcal{L}u(x, t), \quad u(x, 0) = \varphi(x). \quad (6)$$

The operator \mathcal{L} , defined such that

$$(\mathcal{L}\varphi)(x) = \lim_{\tau \rightarrow 0} \frac{\mathbb{E}[\varphi(x(\tau)) \mid x(0) = x] - \varphi(x)}{\tau},$$

for all $\varphi \in \mathcal{S}$, is referred to as the *infinitesimal generator*. As the action of the operators $(\mathcal{P}_t)_{t \geq 0}$ is given as the solution of the differential Equation (6), it is common to use the notation $\mathcal{P}_t := e^{t\mathcal{L}}$. In the case of an Itô diffusion process (5), and with the regularity assumptions on the coefficients \mathbf{a}, \mathbf{B} stated above, it can be shown that the infinitesimal generator takes the form

$$\mathcal{L} = \mathbf{a} \cdot \nabla + \frac{1}{2} \mathbf{B}^T \mathbf{B} : \nabla^2,$$

where “:” denotes the Frobenius product, i.e.,

$$\mathbf{B}^T \mathbf{B} : \nabla^2 = \sum_{i,j=1}^n [\mathbf{B}^T \mathbf{B}]_{i,j} \partial_{x_i} \partial_{x_j},$$

See, e.g., [24] for the case $\mathcal{S} = \mathcal{C}_b^\infty(\Omega_x, \mathbb{R})$, and [25] for extensions from that *core* to more general test function sets such as the ones considered below.

It is often convenient to adopt a dual perspective by considering the evolution of the density of the law μ_t of $\mathbf{x}(t)$ in time. Let $\mu_t(d\mathbf{x}) = \rho(\mathbf{x}, t)d\mathbf{x}$, for $t \geq 0$, in the sense of distributions. The corresponding semigroup $(\mathcal{P}_t^\dagger)_{t \geq 0}$ of *transfer operators* is then naturally defined as the formal adjoint operators of \mathcal{P}_t , i.e.,

$$\int_{\Omega_x} (\mathcal{P}_t \varphi)(\mathbf{x}) \rho_0(\mathbf{x}) d\mathbf{x} = \int_{\Omega_x} \varphi(\mathbf{x}) \mathcal{P}_t^\dagger \rho_0(\mathbf{x}) d\mathbf{x},$$

assuming $\mathbf{x}(0) \sim \mu_0$.

Similarly as for its adjoint, the action of \mathcal{P}_t^\dagger corresponds to the solution of an initial value problem, which in this case is known as the *Fokker-Planck* equation. More specifically,

$$\frac{\partial}{\partial t} \rho(\mathbf{x}, t) = \mathcal{L}^\dagger \rho(\mathbf{x}, t), \quad \rho(\mathbf{x}, 0) = \rho_0(\mathbf{x}), \quad (7)$$

where $\rho(\cdot, t)$ denotes the probability density of μ_t and \mathcal{L}^\dagger —the *Fokker-Planck operator*—can be shown to correspond to the L^2 -adjoint of the infinitesimal generator \mathcal{L} , i.e.,

$$\mathcal{L}^\dagger \rho = -\nabla \cdot (\mathbf{a} \rho) + \nabla^2 : \left(\frac{1}{2} \mathbf{B}^T \mathbf{B} \rho \right),$$

thus

$$e^{t\mathcal{L}^\dagger} = \mathcal{P}_t^\dagger.$$

2.2. Hypocoellipticity and Existence of a Smooth Transition Kernel

Note that (7) is in general to be interpreted in a weak sense as, up to this point, we have not made any assumptions on the regularity of ρ . However, within the scope of this article it is sufficient to consider the case where the differential operator $\partial_t - \mathcal{L}^\dagger$ is *hypoelliptic*. A differential operator A is said to be hypoelliptic, if for any g solving the differential equation $Ag = f$, it follows that g is of higher regularity than f in the sense that

$$f \in H_s^{\text{loc}} \Rightarrow g \in H_{s+\epsilon}^{\text{loc}},$$

with $\epsilon > 0$, where H_s^{loc} denotes the local Sobolev space of order $s \in \mathbb{N}$. This means that if $\partial_t - \mathcal{L}^\dagger$ is hypoelliptic, then the solution of (7) is smooth in the sense that $\rho \in \mathcal{C}^\infty(\Omega_x, (0, \infty))$ irrespective of the regularity of ρ_0 . A common way to establish hypoellipticity of a differential operator is via Hörmander's theorem ([26], Theorem 22.2.1, on p. 353):

Theorem 1. Let A be a differential operator of the form

$$A = \mathbf{a}_0 \cdot \nabla + \sum_{i=1}^M (\mathbf{a}_i \cdot \nabla)^\dagger (\mathbf{a}_i \cdot \nabla),$$

where $\mathbf{a}_i, 0 \leq i \leq M$ are C^∞ vector fields in \mathbb{R}^n and † indicates the formal L^2 adjoint. Iteratively define a collection of vector fields by

$$\mathcal{V}_0 = \{\mathbf{a}_i : i \geq 0\}, \quad \mathcal{V}_{k+1} = \mathcal{V}_k \cup \{[\mathbf{v}, \mathbf{a}_i] : \mathbf{v} \in \mathcal{V}_k, 0 \leq i \leq M\}, \quad (8)$$

where

$$[\mathbf{X}, \mathbf{Y}] = (\nabla \mathbf{Y})\mathbf{X} - (\nabla \mathbf{X})\mathbf{Y},$$

denotes the commutator of vector fields $\mathbf{X}, \mathbf{Y} \in C^\infty(\Omega_x, \mathbb{R}^n)$ and $(\nabla \mathbf{X}), (\nabla \mathbf{Y})$ their Jacobian matrices. If

$$\forall \mathbf{x} \in \mathbb{R}^n, \quad \text{lin} \left\{ \mathbf{v}(\mathbf{x}) : \mathbf{v} \in \bigcup_{k \in \mathbb{N}} \mathcal{V}_k \right\} = \mathbb{R}^n, \quad (9)$$

then the operator A is hypoelliptic.

The condition (9) is commonly referred to as Hörmander's condition. In the particular case of $A = \partial_t - \mathcal{L}^\dagger$ one can easily verify that (9) is exactly satisfied if

$$\mathcal{V}_0 = \{\mathbf{B}_i : i \geq 1\}, \quad \mathcal{V}_{k+1} = \mathcal{V}_k \cup \{[\mathbf{v}, \mathbf{B}_i] : \mathbf{v} \in \mathcal{V}_k, 0 \leq i \leq M\}.$$

with $\mathbf{B}_0 = \mathbf{a}$ and \mathbf{B}_i refers to the i -th column of the diffusion tensor \mathbf{B} in (5). This particular version of Hörmander's condition adapted to the parabolic PDE of the form (7) is referred to as the *parabolic Hörmander condition*.

The direct consequence of the parabolic Hörmander condition is the smoothness of the underlying transition kernel describing the evolution of the probability measure associated to the SDE.

2.3. Ergodicity and Convergence in Law

Let now $\mu(d\mathbf{x}) = \rho(\mathbf{x})d\mathbf{x}$ be a probability measure with density $\rho \in C^2(\Omega_x, [0, \infty))$. When we say that μ is an *invariant measure* of the SDE (5) or that the latter preserves the probability measure μ , we mean that the probability density ρ is a stationary solution of the Fokker-Planck equation associated to the SDE, i.e.,

$$\mathcal{L}^\dagger \rho = 0. \quad (10)$$

Note that for (10) to be well posed in a strong sense it is sufficient that \mathcal{L} is hypoelliptic, i.e., the Fokker-Planck operator \mathcal{L}^\dagger satisfies the Hörmander condition. It is also straightforward to see that every convex combination of two invariant measures is again an invariant measure of the respective SDE which by definition means that the set of invariant measures of a particular SDE is convex.

Let for $t > 0$ and some suitable observable φ ,

$$\bar{\varphi}_t := \frac{1}{t} \int_0^t \varphi(\mathbf{x}(s)) ds,$$

denote the finite time average of φ evaluated along the path of the solution \mathbf{x} of (5). Similarly, we write

$$\mu[\varphi] := \int_{\Omega_x} \varphi(\mathbf{x}) \mu(d\mathbf{x}).$$

for the μ -weighted spatial average of φ . The process \mathbf{x} is said to be *ergodic* with respect to the invariant probability measure μ if for all $\varphi \in L^1(\mu)$ and for almost all realizations of the Wiener process W ,

and μ_0 -almost all initial values $x(0)$, trajectory averages coincide with expectations with respect to the measure μ in the asymptotic limit $t \rightarrow \infty$, i.e.,

$$\lim_{t \rightarrow \infty} \bar{\varphi}_t = \mu[\varphi], \text{ a.s.} \quad (11)$$

It can be shown (see [27]) that ergodicity follows if (i) there exists an invariant measure with positive smooth density and (ii) $\partial_t - \mathcal{L}^\dagger$ is *hypoelliptic*.

Assume that $x(t)$ converges in law towards a unique invariant measure μ , i.e.,

$$\lim_{t \rightarrow \infty} \mathbb{E}[\varphi(x(t)) \mid x(0) = x] = \mu[\varphi], \quad (12)$$

for all $\varphi \in \mathcal{C}_b^\infty(\Omega_x, \mathbb{R})$, μ_0 -almost all $x \in \Omega_x$. A common way to characterize the convergence of the expectation $\mathbb{E}[\varphi(x(t)) \mid x(0) = x]$, or, in semi group notation the convergence of $e^{t\mathcal{L}}\varphi(x)$ to the μ -weighted average $\mu[\varphi]$, is via functional decay estimates of the semi-group operators $e^{t\mathcal{L}}$. For this purpose a set of test functions \mathcal{S} is fixed and equipped with a norm $\|\cdot\|_{\mathcal{S}}$ such that $E := (\mathcal{S}, \|\cdot\|_{\mathcal{S}})$ forms a Banach-space. Of particular interest in this context is exponential convergence of $e^{t\mathcal{L}}\varphi$ towards $\mu[\varphi]$ in the respective norm, i.e.,

$$\|e^{t\mathcal{L}}\varphi - \mu[\varphi]\|_{\mathcal{S}} \leq Ce^{-\kappa t} \|\varphi - \mu[\varphi]\|_{\mathcal{S}}, \quad (13)$$

where C, κ are positive constants, the latter corresponding to the spectral gap of the generator \mathcal{L} in the functional space $E_0 = (\mathcal{S}_0, \|\cdot\|_{\mathcal{S}})$, where $\mathcal{S}_0 \subseteq \mathcal{S}$ denotes the subset of test functions with vanishing mean, i.e.,

$$\mathcal{S}_0 = \{\varphi \in \mathcal{S} : \mu[\varphi] = 0\}.$$

Let the operator Π denote the orthogonal projection from \mathcal{S} onto \mathcal{S}_0 , i.e.,

$$\Pi\varphi = \varphi - \mu[\varphi], \quad \varphi \in \mathcal{S}.$$

Denote further by $\|A\|_{\mathcal{B}(E)}$ the operator norm

$$\|A\|_{\mathcal{B}(E)} := \sup_{\substack{\varphi \in E \\ \varphi \neq 0}} \frac{\|A\varphi\|_{\mathcal{S}}}{\|\varphi\|_{\mathcal{S}}},$$

of an operator $A : E \rightarrow E$. Equation (13) implies that $e^{t\mathcal{L}}\Pi$ when considered as an operator on E is bounded in the operator norm $\mathcal{B}(E)$ as

$$\left\| e^{t\mathcal{L}}\Pi \right\|_{\mathcal{B}(E)} \leq Ce^{-\kappa t}. \quad (14)$$

2.4. Finite-Time Averages and the Central Limit Theorem

Ergodicity of an SDE ensures that time averages of infinitely long trajectories almost surely coincide with spatial averages with respect to the target measure. However, ergodicity per se does not address the statistical properties of finite-time averages apart from the convergence of the time average $\bar{\varphi}_t$ as $t \rightarrow \infty$. For practical applications, where for a given unique invariant measure, the aim is to approximate the μ -weighted average $\mu[\varphi]$ of a test function $\varphi \in \mathcal{S}$, it is important that fluctuations of finite-time averages $\bar{\varphi}_t$ (i.e., the Monte Carlo error in the finite-time approximations) around the infinite-time value $\lim_{t \rightarrow \infty} \bar{\varphi}_t = \mu[\varphi]$ can be quantified. For this purpose it typically regarded as necessary that a central limit theorem holds, i.e.,

$$\sqrt{t}(\bar{\varphi}_t - \mu[\varphi]) \sim \mathcal{N}(0, \sigma_\varphi^2), \text{ as } t \rightarrow \infty, \quad (15)$$

where σ_φ^2 is commonly referred to as the asymptotic variance of the observable φ . A sufficient condition for a central limit theorem of the form (15) to hold for the solution x of (5) and the observable φ is that the Poisson equation

$$-\mathcal{L}\Phi = \Pi\varphi, \quad (16)$$

has a solution which belongs to $L^2(\mu)$ (see [28]). The asymptotic variance then takes the form

$$\sigma_\varphi^2 = -2 \int_{\Omega_x} (\mathcal{L}^{-1}\Pi\varphi) \Pi\varphi d\mu.$$

Let $L_0^2(\mu)$ denote the subspace of $L^2(\mu)$ consisting of functions with vanishing mean. Note that for $\Phi \in L^2(\mu)$ it is a priori not clear how to interpret (16) since only under additional regularity assumptions (16) can be interpreted in a weak sense. A common way to make sense of (16) is by deriving bounds for the operator \mathcal{L}^{-1} in $\mathcal{B}(E_0)$ where E_0 is some subspace of $L_0^2(\mu)$. If \mathcal{L}^{-1} is bounded in $\mathcal{B}(E_0)$, this then directly implies $\Phi \in L^2(\mu)$ in (16) for $\varphi \in E_0$. The relationship between the spectral properties of the operator $\mathcal{L}^{-1}\Pi$ and the convergence properties of the solution process x , or more specifically the decay properties of the semi-group operator $e^{t\mathcal{L}}\Pi$ as $t \rightarrow \infty$ can be made more precise via the formal identity

$$\mathcal{L}^{-1}\Pi = - \int_0^\infty e^{t\mathcal{L}}\Pi dt, \quad (17)$$

which follows from

$$- \int_0^\infty \mathcal{L}e^{t\mathcal{L}}\varphi dt = - \int_0^\infty \left(\frac{d}{dt} e^{t\mathcal{L}} \varphi \right) dt = \varphi, \quad (18)$$

for $\varphi \in \{\varphi \in \mathcal{S}_0 : \mathcal{L}\varphi \in \mathcal{S}_0\}$. Using the identity (17), one directly finds that (14) is a sufficient condition for $\mathcal{L}^{-1}\Pi$ to be bounded since

$$\begin{aligned} \|\mathcal{L}^{-1}\Pi\varphi\|_s &= \left\| \int_0^\infty e^{t\mathcal{L}}\Pi\varphi dt \right\|_s \\ &\leq \int_0^\infty \|e^{t\mathcal{L}}\Pi\varphi\|_s dt \\ &\leq C \int_0^\infty e^{-\kappa t} \|\Pi\varphi\|_s dt \\ &\leq \frac{C}{\kappa} \|\Pi\varphi\|_s. \end{aligned} \quad (19)$$

We conclude that the exponential decay (13) implies the central limit theorem (15) for φ in the corresponding function space E_0 . Estimates for $E_0 = H^1(\mu) \cap L_0^2(\mu)$ can be obtained using the framework of hypocoercivity as presented in [29]. In [30] techniques are introduced to show the decay estimate (13) for $E_0 = L_0^2(\mu)$. In this article we will use Lyapunov function-based techniques which allow to show exponential convergence as in (13) in some weighted L^∞ spaces, which we specify in the next section.

2.5. Exponential Convergence in Weighted L^∞ Spaces

Another way of deriving exponential decay estimates for the semi-group $(e^{t\mathcal{L}})_{t \geq 0}$, which are sufficient to establish a central limit theorem for certain observables, is by means of well established Lyapunov techniques. These techniques have been formulated originally for discrete-time Markov processes/Markov chains [31–33] and have been subsequently extended to continuous time solutions processes of SDEs [22,23,34,35]. The function space on which decay estimates are shown in these references are Banach spaces of the form $(L_{\mathcal{K}}^\infty, \|\cdot\|_{L_{\mathcal{K}}^\infty})$, with

$$L_{\mathcal{K}}^\infty := \left\{ \varphi \text{ measurable} \mid \frac{\varphi}{\mathcal{K}} \in L^\infty(\Omega_x) \right\}, \quad (20)$$

where $\mathcal{K} \in \mathcal{C}^2(\Omega_x, [1, \infty))$ is a positive function, and

$$\|\varphi\|_{L_{\mathcal{K}}^{\infty}} := \left\| \frac{\varphi}{\mathcal{K}} \right\|_{L^{\infty}}. \quad (21)$$

Exponential convergence in the sense of

$$\|e^{t\mathcal{L}}\varphi - \mu[\varphi]\|_{L_{\mathcal{K}}^{\infty}} \leq Ce^{-\kappa t} \|\varphi - \mu[\varphi]\|_{L_{\mathcal{K}}^{\infty}}, \quad (22)$$

is typically shown provided that the following two assumptions are satisfied.

Assumption 1 (Infinitesimal Lyapunov condition). *There is a function $\mathcal{K} \in \mathcal{C}^2(\Omega_x, [1, \infty))$ with $\lim_{\|x\| \rightarrow \infty} \mathcal{K}(x) = \infty$, and real numbers $a \in (0, \infty), b \in (0, \infty)$ such that,*

$$\mathcal{L}\mathcal{K} \leq -a\mathcal{K} + b. \quad (23)$$

Assumption 2. *For some $t > 0$ there exists a constant $\eta \in (0, 1)$ and a probability measure ν such that*

$$\inf_{x \in \mathcal{C}} e^{t\mathcal{L}^+} \delta_x(\mathrm{d}y) \geq \eta \nu(\mathrm{d}y)$$

where $\mathcal{C} = \{x \in \Omega_x : \mathcal{K}(x) \leq \mathcal{K}_{\max}\}$ for some $\mathcal{K}_{\max} > 1 + 2b/a$, where a, b are the same constants as in (23).

As outlined in [23,35], Assumption 2 follows if \mathcal{L} satisfies the parabolic Hörmander condition and the SDE (5) is controllable in the sense that there is a $t > 0$ such that for any pair $x_-, x_+ \in \mathcal{C}$, there exists a continuous control $u \in L^1([0, t], \Omega_x)$, such that the solution \tilde{x} of the differential equation

$$\dot{\tilde{x}} = a(\tilde{x}) + B(\tilde{x})u,$$

satisfies $\tilde{x}(0) = x_-$ and $\tilde{x}(t) = x_+$. The following theorem is derived in [22] using results from [33]. Similar results can be found e.g., in [23,36].

Theorem 2. *Suppose that Assumptions 1 and 2 hold. Then, the solution of the SDE (5) admits a unique invariant probability measure μ , such that*

$$\int_{\Omega_x} \mathcal{K} \mathrm{d}\mu < \infty. \quad (24)$$

Moreover, there exist $C > 0$ and $\kappa > 0$ such that (22) holds for all $\varphi \in L_{\mathcal{K}}^{\infty}(\Omega_x)$ and any $t \in [0, \infty)$.

Finally, let us demonstrate how decay estimates in spaces $L_{\mathcal{K}}^{\infty}$ can be used to derive a central limit theorem for certain functions. Let \mathcal{V} be a Lyapunov function such that the conditions for Theorem 2 are satisfied for $\mathcal{K} = \mathcal{V}$. Note that if the conditions of Theorem 2 are also satisfied for \mathcal{V}^2 , then this implies that a central limit theorem holds for all observables $\varphi \in L_{\mathcal{V}}^{\infty}$, since (24) being valid for $\mathcal{K} = \mathcal{V}^2$ implies

$$L_{\mathcal{V}}^{\infty} \subset L^2(\mu).$$

Thus, the inequality (19) for $\mathcal{S} = L_{\mathcal{V}^2}^{\infty}$ again implies that the solution Φ of (16) is contained in $L^2(\mu)$ for $\varphi \in L_{\mathcal{V}}^{\infty}$, so that by [28] indeed a central limit theorem of the form (15) holds for $\varphi \in L_{\mathcal{V}}^{\infty}$. This motivates to show (22) for a wide class of Lyapunov functions. In the next section, we consider the case of the Langevin dynamics with position dependent coefficients (1) and (2) and show (22) for Lyapunov functions being of the form of polynomials of even but arbitrarily high degree.

3. Langevin Dynamics with Configuration-Dependent Diffusion

We now return to the examples from the introduction, specifically Langevin dynamics with space-dependent friction (1) and (2), and show that in case the variable friction tensor Γ is positive definite and the diffusion tensor Σ has full rank, at all points of the phase space, then the system satisfies the conditions described in the previous section for ergodicity and exponential decay estimates. As illustrations, we consider three applications: (i) a particle model with a temperature gradient; (ii) a simple 2-dimensional Langevin diffusion with a non-conservative force term; (iii) an illustrative Langevin dynamics model for of flocking/swarming, as commonly arises in studies of active particle systems.

3.1. Geometric Ergodicity of Langevin Dynamics With Space-Dependent Coefficients

The main theorem of this section generalizes well known results regarding the ergodicity properties and exponential decay estimates of the associated evolution operator for the underdamped Langevin with constant scalar friction and diffusion coefficient, such as presented in [23,35]. Specifically we extend existing results by

1. allowing the systematic force F to be non-conservative,
2. explicitly considering the case of Γ and Σ being matrix-valued functions of q .

In the current case, no assumption is made regarding a fluctuation-dissipation relation or the form of the invariant density. We will see that under certain conditions on the coefficients Γ and Σ and the non-conservative force F , which are detailed below, the proof of the extended ergodicity criterion for Langevin dynamics follows in a straightforward way from the corresponding proof of the constant-coefficient result [35]. These generalization are nonetheless of high practical relevance and allow us to conclude ergodicity for a wide range of relevant modelling applications.

Let, for a square matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$,

$$\sigma(\mathbf{A}) := \{\lambda \in \mathbb{R} : \exists v \in \mathbb{R}^n, v \neq 0, \lambda v = \mathbf{A}v\},$$

denote its spectrum. The following assumption on the spectrum of Γ and $\Sigma\Sigma^T$ ensures the existence of a suitable Lyapunov function in the case these matrices are not constant in q .

Assumption 3.

- (i) The spectrum $\sigma(\Gamma(q))$ is uniformly bounded in $q \in \Omega_q$ from above and away from 0, i.e., there are positive constants $\lambda_{\min}, \lambda_{\max}$, such that

$$\lambda_{\min} \leq \inf_{q \in \Omega_q} \min \sigma(\Gamma(q)),$$

and

$$\sup_{q \in \Omega_q} \max \sigma(\Gamma(q)) \leq \lambda_{\max}.$$

- (ii) The diffusion matrix has full rank, i.e.,

$$\text{rank}(\Sigma(q)) = n,$$

for all $q \in \Omega_q$, and the spectrum of $\Sigma\Sigma^T$ is bounded in the sense that

$$\exists \bar{\sigma} > 0 : \sup_{q \in \Omega_q} \max \sigma(\Sigma(q)\Sigma^T(q)) \leq \bar{\sigma}.$$

Obviously, Assumption 3 is automatically satisfied for $\Omega_q = L\mathbb{T}^n$ as long as the coefficients Γ, Σ are smooth and $\Gamma(q)$ is positive definite and $\Sigma(q)$ has full rank at every point $q \in \Omega_q$. The next

assumption ensures the existence of a suitable Lyapunov function in the case of a non-conservative force. Again, it is trivially satisfied for $\Omega_q = L\mathbb{T}^n$ and $F \in C^\infty(\mathbb{T}^n, \mathbb{R})$.

Assumption 4. *There exists a potential function $U \in C^2(\Omega_q, \mathbb{R})$ with the following properties*

(i) *there exists $G \in \mathbb{R}$ such that*

$$\forall q \in \Omega_q, \langle q, F(q) \rangle \leq -\langle q, \nabla_q U(q) \rangle + G.$$

for all $q \in \Omega_q$.

(ii) *the potential function is bounded from below, i.e., there exists $u_{\min} > -\infty$ such that*

$$\forall q \in \Omega_q, U(q) \geq u_{\min}.$$

(iii) *there exist constants $D, E > 0$ and $F \in \mathbb{R}$ such that*

$$\forall q \in \Omega_q, \langle q, \nabla_q U(q) \rangle \geq DU(q) + E\|q\|_2^2 + F. \quad (25)$$

We point out that if F is a conservative force, then Assumption 4 reduces to the same asymptotic growth criteria commonly assumed in the derivation of geometric ergodicity of Langevin dynamics with constant coefficients Γ and Σ on an unbounded configurational domain, $\Omega_q = \mathbb{R}^n$ (See again e.g., [23,35]).

Theorem 3. *In (1) and (2), let the force, the friction and diffusion tensors be smooth functions, i.e., $F \in C^\infty(\Omega_q, \mathbb{R}^n)$, $\Gamma \in C^\infty(\Omega_q, \mathbb{R}^{n \times n})$, $\Sigma \in C^\infty(\Omega_q, \mathbb{R}^{n \times m})$ such that Assumptions 3 and 4 hold. There is a unique invariant probability measure*

$$\tilde{\mu}(dq, dp) := \tilde{\rho}(q, p) dq dp,$$

with a (F, Γ, Σ) -dependent density $\tilde{\rho} \in C^\infty(\Omega_q \times \mathbb{R}^n, \mathbb{R}_+)$, and for all $l \in \mathbb{N}$ there are constants $\kappa_l, C_l > 0$ so that

$$\|e^{t\mathcal{L}}\varphi - \tilde{\mu}[\varphi]\|_{L_{\mathcal{K}_l}^\infty} \leq C_l e^{-t\kappa_l} \|\varphi - \tilde{\mu}[\varphi]\|_{L_{\mathcal{K}_l}^\infty}, \quad (26)$$

for all $\varphi \in L_{\mathcal{K}_l}^\infty$ and $t \geq 0$, where

$$\mathcal{K}_l(q, p) = (\langle p, p \rangle + 1)^l, \quad (27)$$

in the case of $\Omega_q = L\mathbb{T}^n$ and

$$\mathcal{K}_l(q, p) = \left(cU(q) + a\langle q, q \rangle + 2b\langle q, p \rangle + \frac{c}{2}\langle p, p \rangle \right)^l, \quad (28)$$

with suitably chosen positive constants $a, b, c > 0$ in the case of $\Omega_q = \mathbb{R}^n$. Furthermore, the probability measure $\tilde{\mu}$ is such, that

$$\int_{\Omega_q \times \mathbb{R}^n} \mathcal{K}_l d\tilde{\mu} < \infty.$$

Proof. Lyapunov condition

We first show that Assumption 1 holds for \mathcal{K}_l as defined in (27) and (28) in the respective setups. Note that

$$\mathcal{L} = \mathcal{L}_H + \mathcal{L}_O,$$

with

$$\begin{aligned} \mathcal{L}_H &= p \cdot \nabla_q - \nabla_q U(q) \cdot \nabla_p, \\ \mathcal{L}_O &= -p^T \Gamma(q) \cdot \nabla_p + \frac{1}{2} \Sigma^T(q) \Sigma(q) : \nabla_q^2. \end{aligned}$$

We first show the existence of constants a_l, b_l such that Assumption 1 is satisfied for $a = a_l, b = b_l$ in the case $l = 1$. For $l \geq 2$, the existence of suitable constants a_l, b_l follows inductively.

By Assumption 4(i), it follows that

$$\begin{aligned}\mathcal{L}_H \mathcal{K}_1(q, p) &= 2b\langle q, F(q) \rangle + 2a\langle p, q \rangle + 2b\langle p, q \rangle \\ &\leq -2b\langle q, \nabla_q U(q) \rangle + G + 2a\langle p, q \rangle + 2b\langle p, q \rangle.\end{aligned}$$

Similarly, by Assumption 3(ii), we find

$$\begin{aligned}\mathcal{L}_O \mathcal{K}_1(q, p) &= -2b\langle q, \Gamma(q)p \rangle - c\langle p, \Gamma(q)p \rangle + \text{ctr}(\Sigma^T(q)\Sigma(q)) \\ &\leq -2b\langle q, \Gamma(q)p \rangle - c\langle p, \Gamma(q)p \rangle + cn\lambda_{\max}.\end{aligned}$$

Putting both inequalities together thus yields

$$\begin{aligned}\mathcal{L} \mathcal{K}_1(q, p) &\leq -2b \left(DU(q) + E\|q\|_2^2 + F \right) + 2(a+b)\langle p, q \rangle - 2b\langle q, \Gamma(q)p \rangle - c\langle p, \Gamma(q)p \rangle + cn\lambda_{\max} \\ &= -2bDU(q) - \begin{pmatrix} q \\ p \end{pmatrix}^T \mathcal{M} \begin{pmatrix} q \\ p \end{pmatrix} - 2bF + cn\lambda_{\max},\end{aligned}$$

with

$$\mathcal{M} = \begin{pmatrix} EbI_n & b\Gamma(q) - aI_n \\ b\Gamma(q) - aI_n & -2bI_n + c\Gamma(q) \end{pmatrix},$$

where we used Assumption 4(iii) in the inequality. The existence of suitable constants a and b such that Assumption 1 is satisfied for the case $\mathcal{K} = \mathcal{K}_1$, follows directly if the matrix \mathcal{M} is positive definite. The positive definiteness of the block matrix \mathcal{M} is implied if (See e.g., [37]), the matrices $-2bI_n + c\Gamma(q)$ and the Schur complement

$$EbI_n - (b\Gamma(q) - aI_n)^T (-2bI_n + c\Gamma(q))^{-1} (b\Gamma(q) - aI_n)$$

are both positive definite. Indeed, since the spectrum of Γ is uniformly bounded on $\Omega_q = \mathbb{R}^n$ according to Assumption 3(i) the positive definiteness of both these matrices can be ensured by choosing a, b sufficiently small and c sufficiently large.

Minorization condition

A simple calculation shows that $\Sigma(q)$ having rank n for all $q \in \Omega_q$ immediately implies that the SDE (1) and (2) satisfies the parabolic Hörmander condition. It therefore only remains to show that for arbitrary $T > 0$ and $(q^-, p^-), (q^+, p^+) \in \mathbb{R}^{2n}$, the control problem

$$\begin{aligned}\frac{dq}{dt} &= p, \\ \frac{dp}{dt} &= F(q) - \Gamma(q)p + \Sigma(q)u, \\ \text{subject to } (q(0), p(0)) &= (q^-, p^-), (q(T), p(T)) = (q^+, p^+),\end{aligned}\tag{29}$$

has a continuous solution $u \in L^1([0, T], \mathbb{R}^m)$. It is easy to verify that there exists a smooth path $\tilde{q} \in C^2([0, T], \mathbb{R}^n)$ such that

$$(\tilde{q}(0), \dot{\tilde{q}}(0)) = (q^-, p^-), (\tilde{q}(T), \dot{\tilde{q}}(T)) = (q^+, p^+).$$

Rewrite (29) as a second order differential equation:

$$\ddot{q} = -\nabla_q U(q) - \Gamma(q)\dot{q} + \Sigma(q)u.$$

Since $\text{rank}(\Sigma(q)) = n$ for all $q \in \Omega_q$, there exists $\Sigma^g \in \mathcal{C}^\infty(\Omega_q, \mathbb{R}^{m \times n})$ such that

$$\Sigma(q)\Sigma^g(q)p = p,$$

for all $q, p \in \mathbb{R}^n$, thus,

$$u(t) = \Sigma^g(\tilde{q}(t)) [\ddot{\tilde{q}} + \nabla_q U(\tilde{q}(t)) + \Gamma(\tilde{q}(t))\dot{\tilde{q}}(t)] \quad (30)$$

is a solution of (29). \square

Note that the case where Γ and Σ are constant was already studied in Mattingly et al. [35] and the proof of the above theorem resembles the structure of the proof therein.

In what follows we provide three examples of models whose ergodic properties can be characterised by the above Theorem 3.

3.2. Single-Particle System with Non-Conservative Force

As an example of a system with a non-conservative force term, which satisfies the condition of the above theorem we consider a Langevin equation of the form (1) and (2) where the force term F is of the form

$$F(q) = -\nabla_q U(q) + \alpha J q, \quad (31)$$

where $J \neq 0$ is a skew-symmetric matrix and $\alpha \in \mathbb{R}$. For $\alpha \neq 0$ the additional term $\alpha J q$ obviously does not correspond to the negative gradient of a smooth potential energy function, thus in this case the force (31) is indeed non-conservative. It is easy to verify that U satisfying the growth conditions (ii) and (iii) in Assumption 4 and

$$U(q) \in O(\|q\|^{2+\epsilon}),$$

with $\epsilon > 0$ as $\|q\| \rightarrow \infty$ implies (i) in Assumption 4. Therefore, as long as the remaining conditions on the coefficients Γ and Σ in Theorem 3 are satisfied, it follows from the same theorem that the respective (non-equilibrium) dynamics possesses an invariant measure, $\mu_\alpha(dq, dp) = \rho_\alpha(q, p)dq dp$, to which it converges exponentially fast in L^∞_K as specified above.

3.3. Multi-Particle Systems

In the remainder of this section we consider the application of Theorem 3 to two different types of particle systems, which can be seen as instances of the underdamped Langevin equation with non-constant coefficients. With some abuse of notation whenever a particle model is considered, let $q_i \in L\mathbb{T}^d, i = 1, \dots, N$ where $L > 0$ and $p_i \in \mathbb{R}^d, i = 1, \dots, N$ denote the position and momentum vectors of the particles, respectively. If, on the other hand, we want to refer to the i -th entry of the vector v where $v = q$ or $v = p$, we write $\Pi_i v$, where $\Pi_i : \mathbb{R}^n \rightarrow \mathbb{R}$ with $i \in \{1, 2, \dots, n\}, n = Nd$, denotes the operator which selects the i -th cartesian coordinate, i.e., $\Pi_i x = e_i \cdot x$. Furthermore, if not stated otherwise, we will assume, that the force is conservative and corresponds to the gradient of a potential function $U(q)$, which is composed as the sum of smooth pair potentials $\tilde{U}_{i,j} \in \mathcal{C}^\infty(\mathbb{R}, \mathbb{R}), 1 \leq j < i \leq N$, i.e.,

$$U(q) = \sum_{i=1}^N \sum_{j=1}^{i-1} \tilde{U}_{i,j}(\|q_i - q_j\|). \quad (32)$$

3.4. Particle System with Temperature Gradient

As a first application we consider an N -particle system with periodic boundary condition where the particles are coupled to a heat bath whose temperature varies depending on the position of the particle within the periodic simulation box. The system we consider in the following is of the form

$$\begin{aligned}\dot{q}_i &= -p_i, \\ \dot{p}_i &= -\nabla_{q_i} U(q) - \gamma p_i + \sqrt{2\gamma(\beta(q_i))}^{-1} \dot{W}_i,\end{aligned}\tag{33}$$

for $i = 1, \dots, N$, where W_i , $i = 1, \dots, N$, are independent white noise processes in \mathbb{R}^d . We further assume that the friction coefficient $\gamma > 0$ is a positive constant and β a smooth positive function on $L\mathbb{T}^d$, i.e., $\beta \in C^\infty(L\mathbb{T}^d, \mathbb{R}_+)$. In light of (29) this corresponds to a constant diagonal friction matrix

$$\Gamma(q) = \gamma I_n,$$

and diffusion tensor of the form

$$\Sigma(q) = \sqrt{2\gamma} \beta^{-\frac{1}{2}}(q) \otimes I_d,$$

where

$$\beta^{-\frac{1}{2}}(q) := \text{diag} \left(\sqrt{\beta(q_1)}, \dots, \sqrt{\beta(q_N)} \right)^{-1}.$$

The matrices Γ and Σ are clearly positive definite and invertible, hence by Theorem 3, there exists a unique invariant measure $\mu_{\gamma, \beta}$ to which the solution of (33) converges (exponentially fast) in law. Note that, even for this relatively simple generalization of the standard underdamped Langevin equation, the form of the invariant measure $\mu_{\gamma, \beta}$ depends nontrivially on γ and β , and one can in general not find an analytic solution of the corresponding stationary Fokker-Planck equation. As long as β^{-1} is bounded from above and away from 0, and the potential U is modified so that Assumption 4 is satisfied (e.g., by adding a confining potential to U), it is easy to see that Theorem 3 applies also for the case $\Omega_q = \mathbb{R}^n$.

3.5. Stochastic Cucker-Smale Model

As a second application of Theorem 3, we consider a stochastic model for flocking which is a variant of the model presented in [4]. In fact, the primary model considered in that paper replaced the deterministic Cucker-Smale system [38], in which a collection of active particles interact with a configuration-dependent friction, by one in which the particles were additionally perturbed by bounded noise. An SDE model was presented there without analysis. A subsequent paper [5] provided numerical evidence for flocking states. In these papers, the SDE approach consists of a Langevin-type model with coordinate-dependent friction and additive (typically uniform) white noise, that is they take the form

$$\frac{dq}{dt} = M^{-1}p,\tag{34}$$

$$\frac{dp}{dt} = -\nabla_q U(q) - \Gamma(q)p + \sigma I_{Nd} \dot{W}(t).\tag{35}$$

For a system of particles moving in one dimension, one has a friction matrix $\Gamma(q) = [\gamma_{ij}(q)]_{1 \leq i, j \leq N} \otimes I_d$, where

$$\gamma_{ij}(q) = -\psi(\|q_j - q_i\|), \quad i \neq j\tag{36}$$

and

$$\gamma_{ii} = -\sum_{j \neq i} \gamma_{ij},\tag{37}$$

for some given scalar kernel function $\psi(r) \geq 0$. That is, the matrix $\Gamma(q)$ is a (weighted) graph Laplacian which reflects the interaction structure of the problem. If $\psi(r) > 0$, this interaction graph is complete.

The inclusion of conservative forces derived from a potential energy function U is an addition to the models mentioned above and allows to incorporate direct attraction and repulsion effects.

Note that since $[\gamma_{ij}]_{1 \leq i, j \leq N}$ is of the form of a graph Laplacian, it is positive semi-definite with at least one eigenvalue being singular. If we assume that $\psi(r) > 0$, then it follows that $[\gamma_{ij}]_{1 \leq i, j \leq N}$ possesses exactly one singular eigenvector, $1_N \in \mathbb{R}^N$, i.e., the vector whose entries all are equal to 1. Consequently, $\Gamma(q)$, has exactly d singular eigenvalues, each of them being of the form $\mathbf{u}_l = 1_N \otimes \mathbf{e}_l$, where \mathbf{e}_l denotes the l -th canonical vector in \mathbb{R}^d . These singular eigenvectors can be associated with the mean momentum

$$\bar{\mathbf{p}} = N^{-1} \sum_{i=1}^N \mathbf{p}_i,$$

of the collection of particles via the relation

$$\bar{\mathbf{p}}_l = N^{-1} \mathbf{u}_l^T \mathbf{p},$$

for $l \in \{1, \dots, d\}$. There are several issues regarding the model (34) and (35). Most importantly, while the diffusion matrix in (35) has full rank, the friction tensor $\Gamma(q)$ is only of rank $(N-1)d$, which means that in directions of the singular vectors \mathbf{u}_l , $1 \leq l \leq d$ there is no dissipation and therefore the kinetic energy of the system will be unbounded as $t \rightarrow \infty$, and hence one cannot expect (34) and (35) to possess an invariant measure. More specifically, if U is composed of pair potentials one can show that the mean momentum $\bar{\mathbf{p}}$ is a Brownian motion,

$$\dot{\bar{\mathbf{p}}} = [\mathbf{u}_l]_{1 \leq l \leq d}^T \dot{\mathbf{W}}.$$

3.5.1. Regularized Stochastic Cucker-Smale Dynamics

A simple fix to the model (34) and (35), which ensures ergodicity of the dynamics, is to add additional dissipation, which is uniform in each component of \mathbf{p} . This can be achieved by replacing the friction tensor in (35) by $\Gamma_\epsilon(q)$, which for $\epsilon > 0$ is defined such that

$$\Gamma_\epsilon(q) = \Gamma(q) + \epsilon \mathbf{I}_{Nd}.$$

It follows directly from the Gershgorin circle theorem that $\Gamma_\epsilon(q)$ is positive definite for all $q \in \Omega_q$ and any choice of $\epsilon > 0$.

3.5.2. Modified Stochastic Cucker-Smale Dynamics

While the above described regularized stochastic Cucker-Smale dynamics is a valid extension of the original model which ensures that the conditions of Theorem 3 are satisfied and hence geometric ergodicity for (34) and (35) holds, the form of the corresponding invariant measure does depend in a non-trivial way on Γ, σ and ϵ and unless Γ is constant in q , one cannot easily find a closed form of the invariant measure. We therefore propose another novel modification of (34) and (35), which is geometrically ergodic with an invariant measure of closed form. We construct this model as a superposition, i.e., $\mathbf{x} = \mathbf{x}_\perp + \mathbf{x}_\parallel$, of two independent stochastic processes, $\mathbf{x}_\perp = (q_\perp, p_\perp)$ and $\mathbf{x}_\parallel = (q_\parallel, p_\parallel)$ taking both values in $\Omega_x \times \mathbb{R}^n$, respectively. We construct these processes such that the parametrization of the former process determines the statistical properties of the stochastic inter-particle interactions and the parametrization of the latter process the collective motion of the flock, i.e., the diffusive properties of the centre of mass. More specifically, we construct \mathbf{x}_\perp as the solution of the SDE

$$\begin{aligned} \frac{dq_\perp}{dt} &= p_\perp, \\ \frac{dp_\perp}{dt} &= -\nabla_q U(q_\perp) - \gamma_\perp \Gamma_\perp(q_\perp) p_\perp + \sqrt{2\gamma_\perp T_\perp} \Sigma_\perp(q_\perp) \dot{\mathbf{W}}(t). \end{aligned} \quad (38)$$

and $\mathbf{x}_{||}$ as the solution of

$$\begin{aligned}\frac{d\mathbf{q}_{||}}{dt} &= \mathbf{p}_{||}, \\ \frac{d\mathbf{p}_{||}}{dt} &= -\gamma_{||}\mathbf{\Gamma}_{||}\mathbf{p}_{||} + \sqrt{2\gamma_{||}T_{||}}\mathbf{\Sigma}_{||}\dot{\mathbf{W}}(t).\end{aligned}\quad (39)$$

where $\gamma_{\perp}, T_{\perp}, \gamma_{||}, T_{||} \geq 0$ and \mathbf{W} as specified in (3) and (4). We refer to (38) as the equation of the *peculiar dynamics* and to (39) as the equation of the *consensus dynamics*. This naming is motivated by the following choices of the respective friction tensor and diffusion matrix.

- Peculiar dynamics:

As in (34) and (35), we choose

$$\mathbf{\Gamma}_{\perp}(\mathbf{q}) = [\gamma_{ij}(\mathbf{q})]_{1 \leq i,j \leq N} \otimes \mathbf{I}_d, \quad (40)$$

where γ_{ij} are defined as above in (36) and (37). By choosing the diffusion matrix $\mathbf{\Sigma}_{\perp}$ such that

$$\mathbf{\Gamma}_{\perp}(\mathbf{q}) = \mathbf{\Sigma}_{\perp}(\mathbf{q})\mathbf{\Sigma}_{\perp}^T(\mathbf{q}),$$

we ensure that the total momentum in each dimension of the physical domain remains constant, i.e.,

$$\frac{d}{dt} \sum_{j=1}^N \mathbf{p}_{\perp j} = \mathbf{0}. \quad (41)$$

This follows directly from the fact, that

$$\mathbf{u}_l \in \ker(\mathbf{\Gamma}_{\perp}(\mathbf{q})) = \ker(\mathbf{\Sigma}_{\perp}(\mathbf{q})), \quad l \in \{1, \dots, d\}.$$

as well as

$$\mathbf{u}_l \cdot \nabla_{\mathbf{q}} U(\mathbf{q}_{\perp}) = 0, \quad (42)$$

(since U is composed of pair potentials) and therefore by (38),

$$\mathbf{u}_l \cdot \frac{d}{dt} \mathbf{p}_{\perp} = 0, \quad l \in \{0, \dots, d\},$$

hence

$$\frac{d}{dt} \sum_{j=1}^N \mathbf{p}_{\perp j} = \sum_{l=1}^d \mathbf{e}_l \mathbf{u}_l \cdot \frac{d}{dt} \mathbf{p}_{\perp} = \mathbf{0}. \quad (43)$$

- Consensus dynamics:

We construct the matrix $\mathbf{\Gamma}_{||}$ such that the difference in the momenta of all particle pairs remains constant under the dynamics (39), i.e.,

$$\frac{d}{dt} (\mathbf{p}_{||j} - \mathbf{p}_{||i}) = \mathbf{0},$$

for all $1 \leq i, j \leq N$. We achieve that by choosing $\mathbf{\Gamma}_{||}$ as

$$\mathbf{\Gamma}_{||} = \mathbf{I}_N \otimes \mathbf{I}_d = [(i+j) \bmod d]_{1 \leq i,j \leq Nd}, \quad (44)$$

and $\mathbf{\Sigma}_{||}$ such that

$$\mathbf{\Gamma}_{||} = \mathbf{\Sigma}_{||}\mathbf{\Sigma}_{||}^T,$$

i.e.,

$$\mathbf{\Sigma}_{||} = N^{-1/2} \mathbf{I}_N \otimes \mathbf{I}_d.$$

- Combined dynamics:

We first observe that although the processes x_{\perp} and x_{\parallel} are driven by the same Wiener process W they are indeed independent. This follows since the column vectors of $\Gamma_{\perp}(q)$ are orthogonal to the column vectors of Γ_{\parallel} in the sense that

$$\ker(\Gamma_{\perp}(q)) = \text{span}(\Gamma_{\parallel}),$$

i.e.,

$$\Gamma_{\perp}(q)\Gamma_{\parallel}^T = \Gamma_{\parallel}\Gamma_{\perp}^T(q) = 0, \quad (45)$$

for all $q \in \Omega_q$, and hence also

$$\Sigma_{\perp}(q)\Sigma_{\parallel}^T = \Sigma_{\parallel}\Sigma_{\perp}^T(q) = 0, \quad (46)$$

so that $\Sigma_{\perp}(q)W$ and $\Sigma_{\parallel}W$ are independent processes, which implies that also the solution processes of the respective SDEs are independent. Moreover, since U is composed of pair-potentials, we have

$$U(q_{\perp} + q_{\parallel}) = U(q_{\perp}). \quad (47)$$

Using (45)–(47) it directly follows that $x = x_{\perp} + x_{\parallel}$ can be identified with the solution of (1) and (2) with

$$\Gamma(q) = \gamma_{\perp}\Gamma_{\perp}(q) + \gamma_{\parallel}\Gamma_{\parallel}, \quad \gamma_{\perp} \geq 0, \gamma_{\parallel} \geq 0, \quad (48)$$

with

$$\Gamma_{\perp}(q) = [\gamma_{ij}(q)]_{1 \leq i,j \leq N} \otimes \mathbf{I}_d, \quad (49)$$

and

$$\Sigma(q) = \sqrt{2\gamma_{\perp}T_{\perp}}\Sigma_{\perp}(q) + \sqrt{2\gamma_{\parallel}T_{\parallel}}\Sigma_{\parallel}, \quad T_{\perp} > 0, T_{\parallel} > 0. \quad (50)$$

or more explicitly with the solution of

$$\begin{aligned} \frac{dq}{dt} &= p, \\ \frac{dp}{dt} &= -\nabla_q U(q) - \gamma_{\perp}\Gamma_{\perp}(q)p - \gamma_{\parallel}\Gamma_{\parallel}p + \left[\sqrt{2\gamma_{\perp}T_{\perp}}\Sigma_{\perp}(q) + \sqrt{2\gamma_{\parallel}T_{\parallel}}\Sigma_{\parallel} \right] \dot{W}(t). \end{aligned} \quad (51)$$

Remark 1. We note that the above choice of $\Gamma_{\perp}(q)$ and $\Sigma_{\perp}(q)$ is very similar to the friction tensor and diffusion tensor in dissipative particle dynamics. In fact (38) would exactly correspond to a DPD system, if instead of (40) one constructs the friction tensor such that dissipation is aligned with the relative orientation of particle pairs

$$\Gamma_{\perp}(q) = [\gamma_{ij}(q) \hat{q}_{i,j} \otimes \hat{q}_{i,j}]_{1 \leq i,j \leq N}.$$

with

$$\hat{q}_{i,j} = (q_j - q_i) / \|q_j - q_i\|.$$

Proposition 1. Let $\Gamma(q)$ and $\Sigma(q)$ defined as above in (49) and (50), with $T_{\perp} > 0, T_{\parallel} > 0$. Let further $\Omega_q = \mathbb{T}^{N \times d}$ and U be of the form (32), then the SDE (51) possesses an invariant measure $\mu_{T_{\perp}, T_{\parallel}}(dq dp) = \rho_{T_{\perp}, T_{\parallel}}(q, p) dq dp$ of the form

$$\rho_{T_{\perp}, T_{\parallel}}(q, p) = \frac{1}{Z} e^{-[T_{\perp}^{-1}U(q) + \frac{1}{2}p^T C^{-1}p]} \quad (52)$$

with covariance matrix

$$C = \frac{1}{N} [T_{\parallel} \mathbf{1}_N + T_{\perp} L] \otimes \mathbf{I}_d$$

where $\mathbf{1}_N \in \mathbb{R}^{N \times N}$ is the all-ones matrix, i.e., every entry of $\mathbf{1}_N$ is equal to 1. The matrix $\mathbf{L} \in \mathbb{R}^{N \times N}$ with

$$\mathbf{L}_{i,j} = \begin{cases} N-1, & i = j, \\ -1, & i \neq j \end{cases}$$

denotes the graph Laplacian of a fully connected graph. In particular,

$$\mathbb{E}_\mu [(\mathbf{p}_j - \mathbf{p}_i) \otimes (\mathbf{p}_j - \mathbf{p}_i)] = 2T_\perp \mathbf{I}_d \quad (53)$$

where $1 \leq i < j \leq N$, and for $\bar{\mathbf{p}} := N^{-1} \sum_{i=1}^N \mathbf{p}_i \in \mathbb{R}^d$,

$$\mathbb{E}_\mu [\bar{\mathbf{p}} \otimes \bar{\mathbf{p}}] = T_\parallel \mathbf{I}_d \quad (54)$$

where $1 \leq j \leq d$.

Proof. We show that (52) is a stationary solution of the Fokker-Planck equation associated with the SDE (51), i.e.,

$$\mathcal{L}^\dagger \rho_{T_\perp, T_\parallel} = 0, \quad (55)$$

with

$$\mathcal{L}^\dagger = \mathcal{L}_H^\dagger + \gamma_\perp \mathcal{L}_{O_\perp}^\dagger + \gamma_\parallel \mathcal{L}_{O_\parallel},$$

where the action of each of these operators applied to $\rho \in \mathcal{C}^2(\Omega_q, \mathbb{R})$ is given as

$$\begin{aligned} \mathcal{L}_H^\dagger \rho &= [-\mathbf{p} \cdot \nabla_q + \nabla_p \cdot \nabla_q U(q)] \rho, \\ \mathcal{L}_{O_\perp}^\dagger \rho &= \nabla_p \cdot (\Gamma_\perp \mathbf{p} \rho) + \nabla_p^2 : T_\perp \Gamma_\perp(q) \rho, \\ \mathcal{L}_{O_\parallel}^\dagger \rho &= \nabla_p \cdot (\Gamma_\parallel \mathbf{p} \rho) + \nabla_p^2 : T_\parallel \Gamma_\parallel \rho. \end{aligned}$$

Before we show (55), we first note that since

$$\ker(\Gamma_\perp(q)) = \text{span}(\Gamma_\parallel),$$

for all $q \in \Omega_q$, we have

$$\Sigma_\perp(q) \Sigma_\parallel^T = \Sigma_\parallel \Sigma_\perp^T(q) = \mathbf{0},$$

hence

$$\begin{aligned} \Sigma(q) \Sigma^T(q) &= \gamma_\perp T_\perp \Sigma_\perp(q) \Sigma_\perp^T(q) + \gamma_\parallel T_\parallel \Sigma_\parallel \Sigma_\parallel^T \\ &= \gamma_\perp T_\perp \Gamma_\perp(q) + \gamma_\parallel T_\parallel \Gamma_\parallel. \end{aligned} \quad (56)$$

Furthermore,

$$\mathbf{C}^{-1} = N^{-1} \left[T_\parallel^{-1} \mathbf{1}_N + T_\perp^{-1} \mathbf{L} \right] \otimes \mathbf{I}_d,$$

which implies

$$(i) \quad \mathbf{C}^{-1} \mathbf{p} = T_\parallel^{-1} \mathbf{p}, \text{ for } \mathbf{p} \in \text{span}(\Gamma_\parallel),$$

$$(ii) \quad \mathbf{C}^{-1} \mathbf{p} = T_\perp^{-1} \mathbf{p}, \text{ for } \mathbf{p} \in \text{span}(\Gamma_\perp).$$

Finally, since $U(q)$ is composed of pair potential functions, we have

$$U(q) = U([c \mathbf{1}_N + N^{-1} \mathbf{L}] \otimes \mathbf{I}_d q),$$

for all $c \in \mathbb{R}$, hence in particular

$$T_\perp^{-1} \nabla_q U(q) = T_\perp^{-1} \nabla_q U \left(\left[\frac{T_\perp}{T_\parallel} \mathbf{1}_N + N^{-1} \mathbf{L} \right] \otimes \mathbf{I}_d q \right) \quad (57)$$

Given the above identities, we conclude

$$\mathcal{L}_H \rho_{T_\perp, T_\parallel} = \left[T_\perp^{-1} \mathbf{p} \cdot \nabla_q U(\mathbf{q}) - \nabla_q U(\mathbf{q}) \cdot \mathbf{C}^{-1} \mathbf{p} \right] \rho_{T_\perp, T_\parallel} = 0$$

due to (57), and

$$\begin{aligned} \mathcal{L}_{O\perp}^+ \rho_{T_\perp, T_\parallel} &= 0, \\ \mathcal{L}_{O\parallel}^+ \rho_{T_\perp, T_\parallel} &= 0, \end{aligned}$$

due to (ii) and (i), respectively. \square

Corollary 1. Let $\psi(r) > 0$ in the definition (36) of γ_{ij} , furthermore $\gamma_\perp > 0$ and $\gamma_\parallel > 0$. The invariant measure $\mu_{T_\perp, T_\parallel}$ specified in Proposition 1 is unique and the law of (51) converges exponentially fast towards $\mu_{T_\perp, T_\parallel}$ in the sense of (26) with $K_l, l \in \mathbb{N}$ as constructed in the proof of Theorem 3.

4. Numerical Discretization

We next describe the construction of numerical methods for the Langevin system. The discretization schemes described here are based on a general splitting framework as explained in the introduction, adapted for the variable coefficient structure.

Following [15], we break the Langevin system (1) and (2) into three parts: A and B corresponding to the deterministic flow:

$$\mathcal{L}_A f = \mathbf{p} \cdot \mathbf{M}^{-1} \nabla_q f$$

and

$$\mathcal{L}_B f = \mathbf{F} \cdot \nabla_p f,$$

and O, which is associated to the isolated momenta diffusion process defined by an Ornstein-Uhlenbeck type equation, i.e., the stochastic system

$$\dot{q} = 0, \quad \dot{\mathbf{p}} = -\Gamma(\mathbf{q}) \mathbf{p} + \Sigma(\mathbf{q}) \dot{\mathbf{W}}(t). \quad (58)$$

A stochastic splitting method is then obtained by concatenating the (stochastic) flow maps corresponding to the B, A and O part. Although other decompositions can be used for splitting, experience in the constant coefficient case has shown that it is beneficial to maintain the same balance of noise and dissipation of the original equations by keeping these terms together. Among stochastic splitting schemes based on this decomposition, the symmetric integration sequence BAOAB was observed to yield a lower discretization bias for ergodic averages in comparison to other integration sequences requiring only one evaluation of the gradient $\nabla_q U$ [16]; moreover, in the case of constant coefficients, this integration sequence has been shown to yield a superconvergent (4th order) error in configurational (q -dependent) observables [15].

In the case of non-constant coefficients $\Gamma(\mathbf{q}), \Sigma(\mathbf{q})$, exact solution of the O-step can be computationally costly. More specifically, bearing in mind that $\mathbf{q}(t) \equiv \mathbf{q}$ is fixed during the isolated Ornstein-Uhlenbeck process (58), a time-step $\Delta t > 0$ can be written:

$$\mathbf{p}(t + \Delta t) = \mathbf{G}_t \mathbf{p}(t) + \mathbf{S}_t \mathbf{R}, \quad (59)$$

where $\mathbf{R} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_{Nd})$, and

$$\mathbf{G}_t = e^{-\Delta t \Gamma(\mathbf{q})}. \quad (60)$$

The matrix $\mathbf{S}_t \in \mathbb{R}^{Nd \times Nd}$ is related to \mathbf{G}_t as

$$\mathbf{S}_t \mathbf{S}_t^T = \mathbf{C}_t - \mathbf{G}_t \mathbf{C}_t \mathbf{G}_t^T, \quad (61)$$

where C_t is the solution of the Lyapunov equation

$$\Gamma(q)C_t + C_t\Gamma^T(q) = \Sigma(q)\Sigma^T(q). \quad (62)$$

This means that in order to compute an exact solution of the O-part as an (59), first the Lyapunov Equation (62) has to be solved, followed by the computation of the Cholesky decomposition (61) and matrix exponential (60). Each of these operations is without any additional assumptions on the structure of the matrices $\Gamma(q)$ and $\Sigma(q)$ of computational complexity $\mathcal{O}(N^3d^3)$. We circumvent these computations by instead integrating the O-part using a numerical method. We construct a symmetric splitting method based on the decomposition

$$\dot{p} = \underbrace{-\Gamma(q)p}_{=:D} + \underbrace{\Sigma(q)\dot{W}}_{=:F},$$

and use the integration sequence DFD, which results in an update of the form

$$p \leftarrow G^2 p + \sqrt{\Delta t} G \Sigma(q) R, \quad R \sim \mathcal{N}(0, I_n), \quad (63)$$

with

$$G = \exp\left(-\frac{\Delta t}{2}\Gamma(q)\right).$$

In order to avoid large numerical errors induced by this approximation we employ a multiple timestepping approach [17], meaning that instead of executing a single integration step of (63) for a full time step, we repeat for $K \in \mathbb{N}$, the integrations step (63) K -times using a step size $\Delta t/K$.

We refer this variant of the BAOAB integration scheme using a multiple time stepping solution of the O-part as “multiple time stepping BAOAB” (m-BAOAB). We provide the implementation of this method in Algorithm 1. The performance evaluation of schemes such as this can be demanding as there is a trade off between computational costs and accuracy of the approximation of the solution of the Ornstein-Uhlenbeck process, which in turn affects the accuracy of the invariant measure of the numerical method. The parameter K in the m-BAOAB algorithm allows control of the accuracy with which the OU process is resolved. The method was tested in the example of Section 5.3, below, but the detailed numerical analysis of this method will be taken up in a forthcoming article.

Algorithm 1 m-BAOAB

```

1: INPUT  $(q, p), \Gamma, \Sigma, \Delta t, K$ 
2:  $q \leftarrow q + (\Delta t/2)p$ 
3:  $p \leftarrow p + (\Delta t/2)F(q)$ 
4:  $G \leftarrow \exp(-\frac{\Delta t}{2K}\Gamma(q))$ 
5: for  $i = 1$  to  $K$  do
6:    $p \leftarrow G^2 p + G\sqrt{\Delta t/K}\Sigma(q)R_i, R_i \sim \mathcal{N}(0, I_n)$ 
7: end for
8:  $p \leftarrow p + (\Delta t/2)F(q)$ 
9:  $q \leftarrow q + (\Delta t/2)p$ 
10: return  $(q, p)$ 

```

5. Numerical Experiments

In this section we consider three systems that fit the form of (1) and (2). In the first, we consider a system of particles diffusing in a temperature gradient. The second model is a two-dimensional example incorporating a nonconservative stirring force. The third example consists of a particle flocking model of the stochastic Cucker-Smale type, as discussed in Section 3. In all of these examples, the model is sufficiently complicated that we do not possess an analytical solution for the nonequilibrium steady

state. For this reason, we are only able to argue for the correctness of the numerical results based on qualitative features (or from our understanding of an unperturbed equilibrium model, in the second example). We are also able to explore the robustness of the numerical solution obtained with respect to variation of the initial conditions, which illustrates the ergodic property. Ultimately comparisons need to be performed with respect to laboratory experiments which can assess the efficacy of the modelling approach in particular situations.

The simulation of the stochastic Cucker-Smale model required the above introduced m-BAOAB scheme for efficient computation. The second model relies on constant (scalar) friction and diffusion coefficient. Due to the diagonal structure of Γ and Σ in the first model, the additional computational costs incurred for the computation of the coefficients in (59) are minimal.

5.1. Particle System with Temperature Gradient

We first study a simple particle system with a position dependent temperature parameter as described in Section 3.4, comprising $N = 64$ particles on a two dimension torus, i.e., $q_i \in L\mathbb{T}^2$ with $L = 5$, which are heated by a source at the center of the simulation box. The heat source is modeled by choosing the position-dependent heat bath temperature as

$$(\beta(q_i))^{-1} = \Psi(|q_i - c_L|),$$

where $c_L = \frac{1}{2}(L, L)$ is the center point of the simulation box and Ψ is a smooth bump function of the form

$$\Psi(r) = \begin{cases} T_{\min} + (T_{\max} - T_{\min}) \exp(-\frac{1}{1-(r/r_{\max})^2}), & r \leq r_{\max} \\ T_{\min}, & r > r_{\max} \end{cases} \quad (64)$$

The constant $T_{\max} > 0$ corresponds to the maximum heat bath temperature at the center point c_L of the simulation box and $T_{\min} > 0$ describes the heat bath temperature outside of the disk

$$B_{r_{\max}}(c_L) := \{q \in L\mathbb{T}^2 : |q - c_L| \leq r_{\max}\}.$$

The potential function U is modeled as the sum of pairwise potentials, i.e.,

$$U(q) = \sum_{i=1}^{N-1} \sum_{j=1}^{i-1} \omega(r_{i,j})$$

with $r_{i,j} := |q_j - q_i|$ and ω is a simple repulsive soft potential of the form

$$\omega(r_{i,j}) = \begin{cases} \frac{k}{2}(r_{i,j} - c_r)^2, & r_{i,j} < c_r \\ 0, & r_{i,j} \geq c_r. \end{cases}$$

where k and c_r are positive constants. The pair interaction described by ω corresponds to a harmonic spring of stiffness k and rest length c_r . Particle systems involving this type of pair potential are commonly used as benchmark systems in the context of dissipative particle dynamics [39,40]. Due to the isolated jump discontinuity in the second derivative of the potential Theorem 3 does not strictly speaking, apply in this case. Although it would be easy to modify the potential to have any desired level of smoothness, we expect, based on our experience with molecular dynamics problems where similar such issues arise, that the results will be very similar to those obtained with the potential given here. The simulation results reported in the remainder of this section are obtained for a parameterization of the model where $k = 25, c_r = 1$ and $r_{\max} = 1$ with $T_{\min} = 1/10$ and $T_{\max} = 4$. The particle positions and momenta were initialized on an equidistant grid such that $q_i = \frac{L}{N}(\lfloor \frac{i}{N} \rfloor, i \bmod N)$ and $p_i = (0, 0)$, respectively. $N_t = 10^5$ timesteps of stepsize $h = 2 \times 10^{-2}$ were simulated with varying values $\gamma \in \{10^i, i = -1, 0, 1, 2\}$ of the friction coefficient. Define by

$$\beta_{\gamma}^{-1}(r) := \mathbb{E}(p_i^2 \mid |q_i - c_L| = r)$$

the *effective temperature* of the system at distance $r \in [0, L/2)$ from the center c_L of the simulation box. Figure 1B, shows estimates of this function for different values of γ calculated at points $r_i = \Delta_r/2 + i\Delta_r, i = 0, \dots, \lfloor \frac{L}{2\Delta_r} \rfloor$, as

$$\hat{\beta}_{\gamma}^{-1}(r_i) = \frac{c_i}{N_i} \sum_{j=1}^N \sum_{k=1}^{N_i} \mathbb{1}_{[r_i - \Delta_r/2, r_i + \Delta_r/2)} (|q_j^{(k)} - c_L|) \frac{1}{2} \left((p_{j,1}^{(k)})^2 + (p_{j,2}^{(k)})^2 \right), \quad (65)$$

with

$$c_i := \frac{1}{[(r_i + \Delta_r/2)^2 - (r_i - \Delta_r/2)^2]} \left(\sum_{i=1}^{\lfloor \frac{L}{2\Delta_r} \rfloor} \frac{1}{(r_i + \Delta_r/2)^2 - (r_i - \Delta_r/2)^2} \right)^{-1},$$

and

$$N_i = \sum_{j=1}^N \sum_{k=1}^{N_i} \mathbb{1}_{[r_i - \Delta_r/2, r_i + \Delta_r/2)} (|q_j^{(k)} - c_L|).$$

A larger value of γ leads to a tighter local coupling of the particle system with the heat bath and hence for large friction $\gamma = 10^2$, the estimates of the effective temperature coincide very well with the heat bath temperature given by Ψ at the respective distances from the center of the simulation box. For smaller friction values the position dependence of the effective temperature is less apparent and for very small friction values, e.g., $\gamma = 10^{-2}$, it is nearly lost entirely. At the same time the decay of the temperature gradient for smaller friction values leads to an increase in the temperature outside the heat source region. Indication of this effect can be also found in the estimated radial density function (see Figure 1A). For small values of the friction coefficient the increased temperature leads to a stronger fluctuation of inter-particle distances outside the heat source area, and hence minima and maxima in the density of radial distribution are less apparent in comparison to the radial distributions for systems with a high friction value.

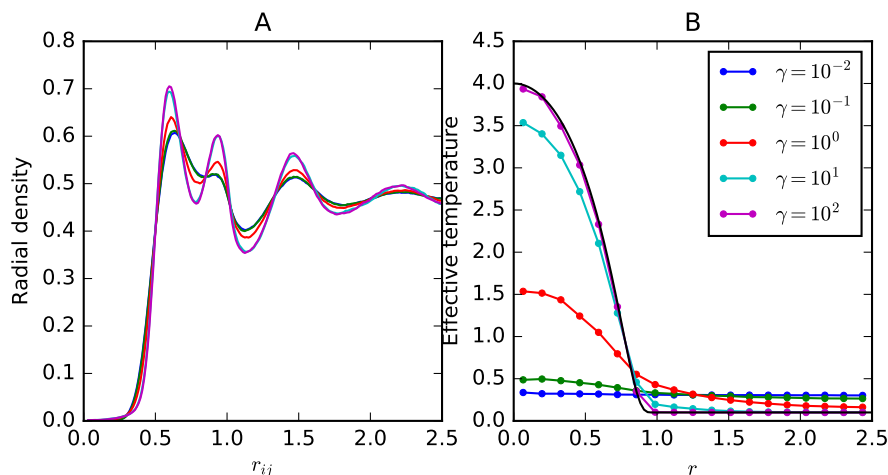


Figure 1. (A) Radial density function for the particle system described in Section 5.1 for different values of the friction coefficient γ ; (B) Distance r from heat source center c_L vs. effective temperature estimated according to (65). The black curve corresponds to the heat bath temperature $\Psi(r)$, with Ψ as defined in (64).

Due to the relatively small size of the heat source area the number of particles in this area is small in comparison to the number of particles outside of the heat source area and hence the shape of the radial density is mainly determined by the statistical properties of the particles outside the heat source

area. The empirical particle densities shown in Figure 2 provide further insight into how the value of the friction coefficient affects the properties of $\mu_{\gamma,\beta}$. For a small friction coefficient, i.e., $\gamma = 10^{-2}$ the particle density is very close to uniform on $L\mathbb{T}^2$. If the heat bath temperature was constant in $L\mathbb{T}^2$ one would expect a uniform density as the potential function U is solely composed of pair potentials and thus translation-invariant in the coordinates of the particle density. More precisely, let $q_{i,1}$ and $q_{i,2}$ denote the first and second coordinate value of the i -th particle, then

$$U(q) = U((q_{i,1} + a_1, q_{i,2} + a_2)_{1 \leq i \leq N}),$$

for all $a_1, a_2 \in \mathbb{R}$. Therefore, the observed uniform particle density for $\gamma = 10^{-2}$ is consistent with the observation that the position dependency of the effective temperature vanishes for this friction value. While for the considered trajectory length/sample size the plot of the invariant for a friction coefficient $\gamma = 10^2$ is very noisy, it still strongly suggests that the translation invariance of the particle density, i.e., the uniformity of the particle density, is broken in this regime. Within the interior of the heat source region particles are distributed approximately uniformly whereas close to the boundary of the heat source region the particle density is increased. Outside the heat source region the particle density is concentrated around positions close to energy minima of the potential energy function U .

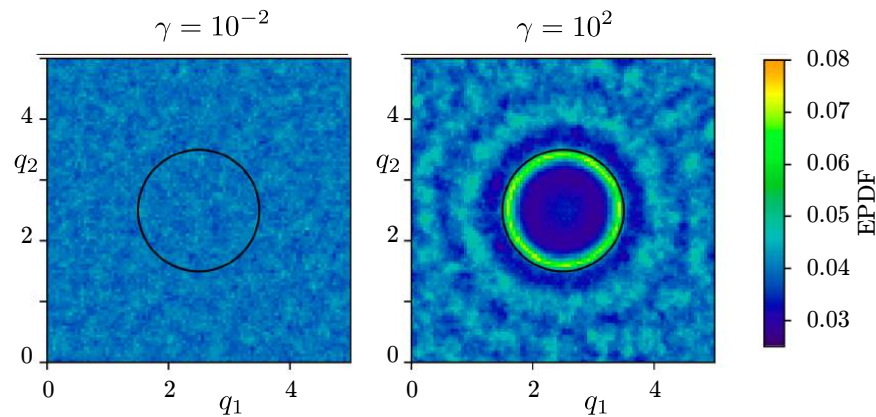


Figure 2. Empirical particle density calculated as a cumulative average over the simulation time, for two values of γ . The area inside the black circle corresponds to the heat source area.

5.2. System with Non-Conservative Force

We next consider an instance of a non-equilibrium system with a non-conservative force of a form as outlined in Section 3.2. More specifically, we let $\Omega_q = \mathbb{R}^n$ with $n = 2$ and let the force F be of the form (31), i.e.,

$$F(q) = -\nabla_q U(q) + \alpha J q,$$

with

$$U(q) = \sum_{i=1}^2 (q_i^2 - 1)^2,$$

and

$$J := \frac{3}{2} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

We furthermore assume that the system is driven by a standard Langevin diffusion, i.e.,

$$\Gamma = \gamma I_n,$$

and

$$\Sigma = \sqrt{2\gamma\beta^{-1}}\mathbf{I}_n,$$

where we choose $\gamma = 1$, and $\beta = 1$. This means, that without the non-conservative force part, i.e., in the case $\alpha = 0$, the system considered resembles to a particle moving in a 4-well potential driven by a standard Langevin equation at equilibrium at unit temperature. The non-conservative force part $\alpha \mathbf{J} \mathbf{q}$, corresponds to a stirring force, which pushes the system radially and in clockwise direction around the origin. The effect of the stirring force can be seen in Figure 3. In the absence of the stirring force (see Figure 3A) the invariant distribution is exactly the canonical distribution, i.e.,

$$\rho_0(\mathbf{q}, \mathbf{p}) \propto \exp\left(-U(\mathbf{q}) - \frac{1}{2}\mathbf{p} \cdot \mathbf{p}\right).$$

The modes of the marginal density in \mathbf{q} are exactly positioned at the energy minima of U at $(\pm 1, \pm 1)$. Estimates of the mean momentum vectorfield

$$\bar{\mathbf{p}}(\mathbf{q}) := \int_{\Omega_q} \mathbf{p} \rho_\alpha(\mathbf{q}, \mathbf{p}) d\mathbf{p}, \quad (66)$$

vanish at each point in the configurational phase space. In the presence of the stirring force (see Figure 3B), the invariant density is rotated slightly and smeared out over the energy barriers. The mean momentum resembles a vector field spiralling clockwise around the origin.

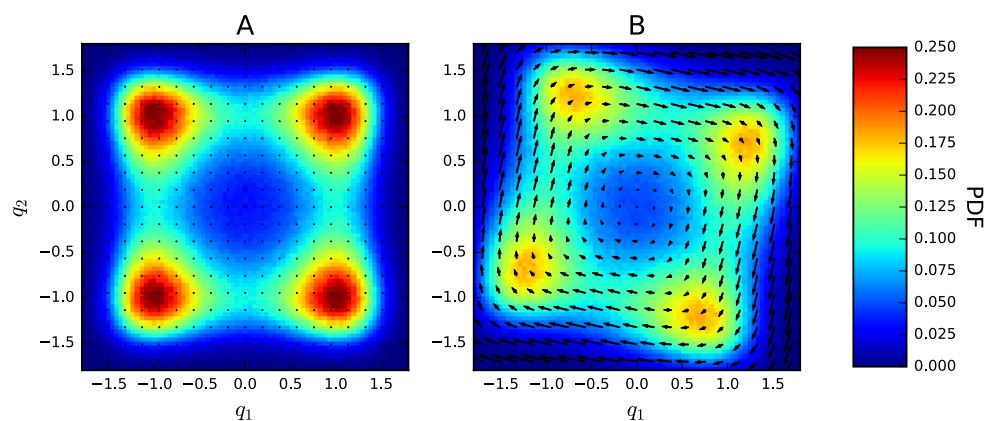


Figure 3. Marginal of \mathbf{q} of the invariant density of the system described in Section 5.2 in the absence ($\alpha = 0$) of the non-conservative force (A), and in the presence ($\alpha = 1$) of the non-conservative force (B). The black arrows correspond to estimates of the mean momentum vector field (66) in the invariant density at the respective points in configurational space.

5.3. Stochastic Cucker-Smale Model

In this section we present simulation results for a modified stochastic Cucker-Smale model as described in Section 3.5.2. In particular, we demonstrate that the peculiar and consensus temperature as well as the consensus diffusion rate can be controlled independently and match in simulation with the analytically derived values, if the friction tensor Γ and the diffusion tensor Σ are suitably parametrized. Apart from these macroscopic quantities, we also demonstrate the effect of different parameter values of $\gamma_\perp, \gamma_\parallel, T_\perp, T_\parallel$ on other macroscopic quantities which in general do not possess a closed form, such as the radial density of particles and peculiar velocity autocorrelation function. Several animations of the particle motion, which effect of parameter choices, may be viewed at the following web location: <https://github.com/MatthiasSachs/StochasticCuckerSmale.git>.

5.3.1. Model Parametrization

If not stated otherwise the results presented in this section were obtained from a system of 64 particles in a periodic simulation box of edge length 5 and dimension 2, i.e., $\Omega_q = L\mathbb{T}^{Nd}$ with $d = 2, L = 5$ and $N = 64$. We assume the particle pair-potentials $\tilde{U}_{i,j}$ in the definition of U to be identical Morse potentials, i.e., for $i \neq j$

$$\tilde{U}_{i,j}(r) = D(1 - e^{-a(r-c_r)})^2, \quad r > 0,$$

with $D = 1, a = 1, c_r = 1/2$. Following [4] we choose the functions ψ in the definition of γ_{ij} in (36) to be of the form

$$\psi(r) = \frac{K}{1 + \|r\|^\alpha},$$

with $K = 1/10$, and $\alpha = 6$.

5.3.2. Independent Control of Peculiar and Consensus Temperature

We first demonstrate that as stated in Proposition 1, we can indeed control the peculiar and the consensus temperature independently, such that these quantities coincide with the values of the model parameters T_{\parallel} and T_{\perp} , respectively, i.e., we show that the identities (53) and (54) are reproduced in simulations. For this purpose we compute an estimate of the peculiar temperature at time step $n \in \mathbb{N}$ as

$$\varphi_{T_{\perp}}(\mathbf{p}^{(k)}) = \frac{1}{N(N-1)d} \sum_{i=1}^{N-1} \sum_{j=i+1}^N \sum_{l=1}^d \left(\mathbf{p}_{j,l}^{(k)} - \mathbf{p}_{i,l}^{(k)} \right)^2 \quad (67)$$

and an estimate of the consensus temperature at time step $k \in \mathbb{N}$ as

$$\varphi_{T_{\parallel}}(\mathbf{p}^{(k)}) = \frac{1}{d} \sum_{l=1}^d \left(\bar{\mathbf{p}}_l^{(k)} \right)^2. \quad (68)$$

In Figure 4 we show the results for a system with $T_{\perp} = 1/2$ and $T_{\parallel} = 5$. We see that the cumulative average of the respective estimates converges, after a short equilibration period, to the target values.

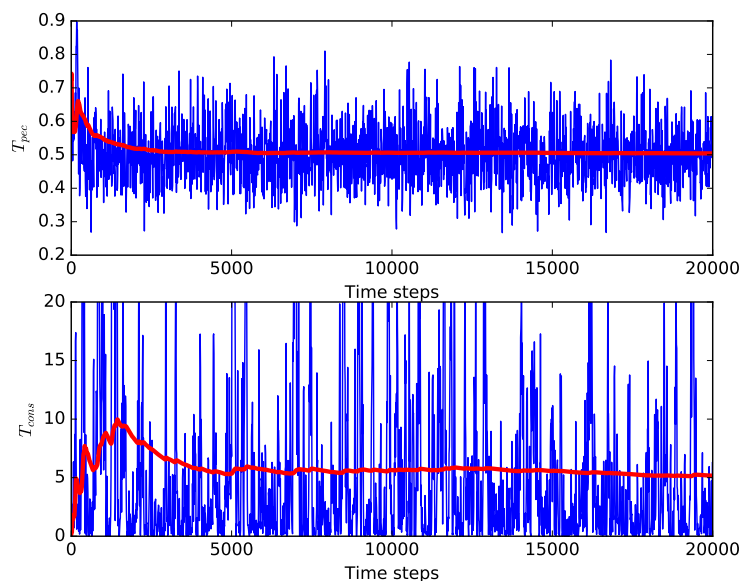


Figure 4. Time vs. observed peculiar temperature (upper figure) and consensus temperature (lower figure). The blue trajectory shows the estimates $\varphi_{T_{\perp}}(\mathbf{p}^{(k)})$ and $\varphi_{T_{\parallel}}(\mathbf{p}^{(k)})$ at timestep k for the peculiar and the consensus temperature, respectively. The respective cumulative averages are shown in red.

5.3.3. Properties of the Flock

We first demonstrate the effect of the parameterization of the peculiar dynamics (the values of the peculiar friction γ_{\perp} and the peculiar temperature T_{\perp}) on the flock size, flock formation, and inter-flock diffusion. While we vary the values of the peculiar temperature parameter and peculiar friction parameter in order to demonstrate the effect of these parameters on the above quantities, we leave the parameter values of the consensus dynamics unchanged in all simulations as $\gamma_{\parallel} = 1$ and $T_{\parallel} = 1$. In the first series of simulations we vary the value of T_{\perp} , while fixing the value of the peculiar friction parameter to $\gamma_{\perp} = 1$. Figure 5 shows the radial density for different values of T_{\perp} . As one would expect, the probability mass is concentrated around the mean rest-length of the Morse-potential for low values of T_{\perp} and the distribution spreads out for higher values of T_{\perp} . If we measure flock size in terms of the mean distance between particles in the flock, this implies that the flock size grows for increased temperature.

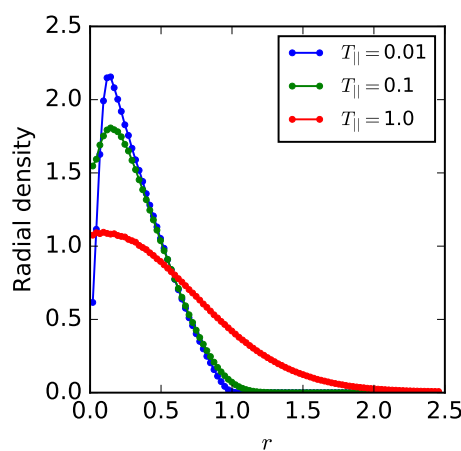


Figure 5. Radial density function for varying values of the peculiar temperature parameter T_{\perp} .

We next explore the effects of the peculiar friction parameter γ_{\perp} . We study the effect of the value of γ_{\perp} on the flock formation using the mean distance between particles,

$$\varphi_{\text{md}}(\mathbf{q}) = \frac{2}{N(N-1)} \sum_{i=1}^{N-1} \sum_{j=i+1}^N \|\mathbf{q}_i - \mathbf{q}_j\|,$$

as a measure of the progression of the flock formation. We initialize the particle position out of equilibrium on a equidistant square grid covering a square with side length $5/2$. Figure 6A shows the time evolution of observed mean distance $\varphi_{\text{md}}(\mathbf{q})$ in simulations using different values of γ_{\perp} . We find that for small value of γ_{\perp} flock formation comes with strong oscillation in the flock size (measured in terms of the mean particle distance), which only slowly decays. With increased values of γ_{\perp} these oscillations are more strongly damped so that the flock size quickly approaches its equilibrium value.

We next explore the effect of the value of the model parameter γ_{\perp} on the mobility of particles within the flock. In order to measure the strength of diffusion within the flock we consider the mean distance of the particle i to the center of mass, i.e.,

$$\varphi_{\text{mc}}^i(\mathbf{q}) = \|\mathbf{q}_i - \bar{\mathbf{q}}\|.$$

The autocorrelation of this observable can be used as a measure of mobility within the flock. Figure 6B, shows the autocorrelation function for φ_{mc}^i calculated from a trajectory in equilibrium, i.e., after the initial transitional flock-formation phase described in the previous paragraph. We can see that the decay of the autocorrelation function becomes slower for increasing values of γ_{\perp} , which indicates that for an increased value of γ_{\perp} the mobility of particles within the flock is reduced.

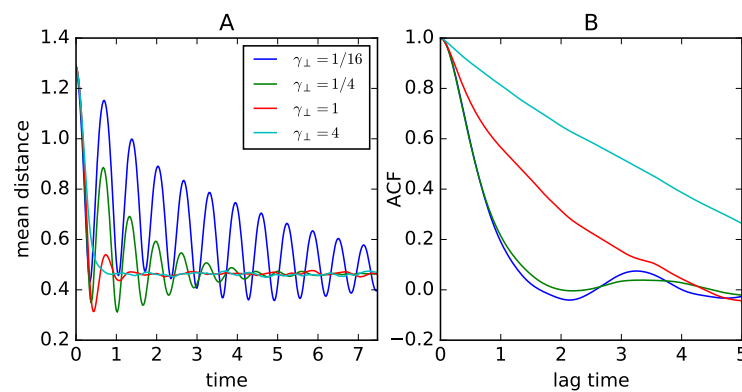


Figure 6. Flock formation and inter-flock diffusion for different values of the peculiar friction parameter γ_{\perp} . **(A)** Time vs. mean distance $\phi_{md}(q)$; **(B)** Lag time vs. time-autocorrelation of the mean-centre distance $\phi_{mc}^i(q)$. The plotted curve is computed as an average over all particles indices $i = 1, \dots, N$.

5.3.4. Collective Motion

We next explore the effects of the values of γ_{\parallel} and T_{\parallel} on the collective motion of the flock, i.e., the diffusive behaviour of the center of mass. Since the motion of the center of mass is described by the *consensus dynamics*, which is driven by an Ornstein-Uhlenbeck process, we find the diffusion constant D of the center of mass (viewed as a free particle in space) to be

$$D = \frac{T_{\parallel}}{N\gamma_{\parallel}}.$$

Table 1 shows the estimated diffusion coefficients for various values of γ_{\parallel} and T_{\parallel} . We find a good match between theoretically predicted values and observed values.

Table 1. Estimates of ND for various values of γ_{\parallel} and T_{\parallel} . $N = 16$ denotes the number of particles and D the diffusion coefficient of the diffusive motion of the center of mass. Values in parentheses show the exact values.

$\gamma_{\parallel}^{-1} \backslash T_{\parallel}$	1	10	100
1	0.987, (1)	9.775, (10)	101.8, (100)
3	3.059, (3)	29.445, (30)	308.1, (300)
9	8.476, (9)	91.994, (90)	861.4, (900)

6. Conclusions

In this article we have provided a general treatment of the convergence of Langevin dynamics to a stationary state, including for systems with configuration-dependent friction and noise, as well as nonconservative forces. We have demonstrated the concepts in applications to systems with temperature gradients and stochastic models of active particle systems. Our approach does not assume the usual fluctuation-dissipation rule, so it can be applied to a wide range of nonequilibrium molecular and particle systems where the form of the stationary distribution is a priori unknown. Future work might look at the use of configuration-dependent memory kernels (within a generalized Langevin equation setting) and proofs of ergodicity for degenerate thermostats, for example the pairwise adaptive thermostats [40], which provide an alternative method for controlling observables.

Acknowledgments: The research of all three authors was supported by the ERC project RULE (grant number 320823). The collaboration was initiated during a residency of the first two authors at the Institut Henri Poincaré and its program on “Stochastic dynamics out of equilibrium.” The work of M. Sachs was further supported by the Statistical and Applied Mathematical Sciences Institute (North Carolina).

Author Contributions: The article was conceived in joint discussions of all three authors during May of 2017. All three authors contributed to the writing. The numerical experiments were proposed and discussed by all three authors but were performed by M.S.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Lancon, P.; Batrouni, G.; Lobry, L.; Ostrowsky, N. Drift without flux: Brownian walker with a space-dependent diffusion coefficient. *Europhys. Lett.* **2001**, *54*, 28–34.
2. Regev, S.; Grønbech-Jensen, N.; Farago, O. Isothermal Langevin dynamics in systems with power-law spatially dependent friction. *Phys. Rev. E* **2016**, *94*, 012116.
3. Becton, M.; Wang, X. Thermal gradients on graphene to drive nanoflake motion. *J. Chem. Theory Comput.* **2014**, *10*, 722–730.
4. Cucker, F.; Mordecki, E. Flocking in noisy environments. *J. Math. Pures Appl.* **2008**, *89*, 278–296.
5. Ha, S.Y.; Lee, K.; Levy, D. Emergence of time-asymptotic flocking in a stochastic Cucker-Smale system. *Commun. Math. Sci.* **2009**, *7*, 453–469.
6. Gachelin, J.; Rousselet, A.; Lindner, A.; Clement, E. Collective motion in an active suspension of *Escherichia coli* bacteria. *New J. Phys.* **2014**, *16*, 025003.
7. Best, R.; Hummer, G. Coordinate-dependent diffusion in protein folding. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 1088–1093.
8. Chen, C.; Liu, S.; Shi, X.Q.; Chaté, H.; Wu, Y. Weak synchronization and large-scale collective oscillation in dense bacterial suspensions. *Nature* **2017**, *542*, 210–214.
9. Hoogerbrugge, P.; Koelman, J. Simulating microscopic hydrodynamic phenomena with dissipative particle dynamics. *Europhys. Lett.* **1992**, *19*, 155–160.
10. Español, P. Dissipative particle dynamics. In *Handbook of Materials Modeling*; Springer: Berlin, Germany, 2005; pp. 2503–2512.
11. Shardlow, T.; Yan, Y. Geometric ergodicity for dissipative particle dynamics. *Stoch. Dyn.* **2006**, *6*, 123–154.
12. Ahn, S.; Ha, S.Y. Stochastic flocking dynamics of the Cucker-Smale model with multiplicative white noises. *J. Math. Phys.* **2010**, *51*, 103301.
13. Ton, T.; Linh, N.; Yagi, A. Flocking and non-flocking behavior in a stochastic Cucker-Smale System. *Anal. Appl.* **2014**, *12*, 63–73.
14. Erban, R.; Haskovec, J.; Sun, Y. A Cucker-Smale model with noise and delay. *SIAM J. Appl. Math.* **2016**, *76*, 535–1557.
15. Leimkuhler, B.; Matthews, C.; Stoltz, G. The computation of averages from equilibrium and nonequilibrium Langevin molecular dynamics. *IMA J. Numer. Anal.* **2016**, *36*, 13–79.
16. Leimkuhler, B.; Matthews, C. Rational construction of numerical methods for stochastic molecular dynamics. *Appl. Math. Res. Exp.* **2013**, *2013*, 34–56.
17. Tuckerman, M.; Berne, B.J.; Martyna, G.J. Reversible multiple time scale molecular dynamics. *J. Chem. Phys.* **1992**, *97*, 1990–2001, doi:10.1063/1.463137.
18. Hummer, G. Position-dependent diffusion coefficients and free energies from Bayesian analysis of equilibrium and replica molecular dynamics simulations. *New J. Phys.* **2005**, *7*, 34.
19. Burada, P.; Schmid, G.; Reguera, D.; Rubi, J.; Hänggi, P. Biased diffusion in confined media: Test of the Fick-Jacobs approximation and validity criteria. *Phys. Rev. E* **2007**, *75*, 051111.
20. Berezhkovskii, A.; Szabo, A. Time scale separation leads to position-dependent diffusion along a slow coordinate. *J. Chem. Phys.* **2011**, *135*, 074108.
21. Wilmer, O.R.; Pedro, C.; Lopez, F. Direct evaluation of the position dependent diffusion coefficient and persistence time from the equilibrium density profile in anisotropic fluids. *J. Chem. Phys.* **2013**, *139*, 074103.
22. Lelièvre, T.; Stoltz, G. Partial differential equations and stochastic methods in molecular dynamics. *Acta Numer.* **2016**, *25*, 681–880.

23. Bellet, L.R. Ergodic properties of Markov processes. In *Open Quantum Systems II*; Springer: Berlin, Germany, 2006; pp. 1–39.
24. Gardiner, C. Handbook of stochastic methods for physics, chemistry and the natural sciences. *Appl. Opt.* **1986**, *25*, 3145.
25. Guionnet, A.; Zegarliński, B. Lectures on logarithmic Sobolev inequalities. In *Séminaire de Probabilités XXXVI*; Springer: Berlin, Germany, 2003; pp. 1–134.
26. Hörmander, L. *The Analysis of Linear Partial Differential Operators. III, Volume 274 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*; Springer, Berlin, Germany, 1985.
27. Kliemann, W. Recurrence and invariant measures for degenerate diffusions. *Ann. Probab.* **1987**, *15*, 690–707.
28. Bhattacharya, R.N. On the functional central limit theorem and the law of the iterated logarithm for Markov processes. *Zeitschrift Für Wahrscheinlichkeitstheorie Und Verwandte Gebiete* **1982**, *60*, 185–201.
29. Villani, C. *Hypocoercivity*; Number 949-951; American Mathematical Society: Providence, RI, USA, 2009.
30. Dolbeault, J.; Mouhot, C.; Schmeiser, C. Hypocoercivity for linear kinetic equations conserving mass. *Trans. Am. Math. Soc.* **2015**, *367*, 3807–3828.
31. Harris, T.E. The existence of stationary measures for certain Markov processes. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, Volume 2: Contributions to Probability Theory*; University of California Press: Oakland, CA, USA, 1956; pp. 113–124.
32. Meyn, S.P.; Tweedie, R.L. Stability of Markovian processes I: Criteria for discrete-time chains. *Adv. Appl. Probab.* **1992**, *24*, 542–574.
33. Hairer, M.; Mattingly, J.C. Yet another look at Harris’ ergodic theorem for Markov chains. In *Seminar on Stochastic Analysis, Random Fields and Applications VI*; Springer: Berlin, Germany, 2011; Volume 63, pp. 109–117.
34. Meyn, S.P.; Tweedie, R.L. Stability of Markovian processes III: Foster–Lyapunov criteria for continuous-time processes. *Adv. Appl. Probab.* **1993**, *25*, 518–548.
35. Mattingly, J.C.; Stuart, A.M.; Higham, D.J. Ergodicity for SDEs and approximations: Locally Lipschitz vector fields and degenerate noise. *Stoch. Process. Their Appl.* **2002**, *101*, 185–232.
36. Meyn, S.P.; Tweedie, R.L. *Markov Chains and Stochastic Stability*; Springer Science & Business Media: Berlin, Germany, 2012.
37. Zhang, F. *The Schur Complement and Its Applications*; Springer Science & Business Media: Berlin, Germany, 2006; Volume 4.
38. Cucker, F.; Smale, S. Emergent behavior in flocks. *IEEE Trans. Autom. Control* **2007**, *52*, 852–862.
39. Groot, R.D.; Warren, P.B. Dissipative particle dynamics: Bridging the gap between atomistic and mesoscopic simulation. *J. Chem. Phys.* **1997**, *107*, 4423–4435.
40. Leimkuhler, B.; Shang, X. Pairwise adaptive thermostats for improved accuracy and stability in dissipative particle dynamics. *J. Comput. Phys.* **2016**, *324*, 174–193.

